Уральский
федеральный
университет
имени первого Президента
России Б. Н. Ельцина
Институт радиоэлектроники
и информационных
технологий — РТФ

М. А. АЛЬШАНСКИЙ

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Учебное пособие

Министерство науки и высшего образования
Российской Федерации

Уральский федеральный университет
имени первого Президента России Б. Н. Ельцина

М. А. Альшанский

Теория вероятностей и математическая статистика

Учебное пособие

Рекомендовано методическим советом
Уральского федерального университета
для студентов вуза, обучающихся
по направлениям подготовки
11.03.01 — Радиотехника,
11.03.02 — Инфокоммуникационные технологии
и системы связи

Екатеринбург
Издательство Уральского университета
2021

УДК 519.2(075.8)
ББК 22.17я73
А56

Р е ц е н з е н т ы:
канд. физ.-мат. наук, доцент, старший научный сотрудник Института математики и механики УрО РАН *В. Л. Розенберг*;
кафедра высшей математики и методики обучения математике ФГБОУ ВО «Уральский государственный педагогический университет» (завкафедрой, д-р физ.-мат. наук *В. Ю. Бодряков*)

Альшанский, Максим Алексеевич.
А56 Теория вероятностей и математическая статистика : учебное пособие / М. А. Альшанский ; М-во науки и высшего образования РФ. — Екатеринбург : Изд-во Урал. ун-та, 2021. — 224 с.

ISBN 978-5-7996-3385-1

Учебное пособие содержит материал курса теории вероятностей и математической статистики, подготовленный для обучения студентов ИРИТ-РтФ Уральского федерального университета. Изложение материала дает читателю возможность получить представление о фундаментальных основах предмета без чрезмерного погружения в детали теории меры и интеграла Лебега и опирается на стандартные курсы высшей математики для втузов. Учебное пособие предназначено для студентов инженерных специальностей УрФУ.

УДК 519.2(075.8)
ББК 22.17я73

ISBN 978-5-7996-3385-1

Оглавление

Введение

Настоящее учебное пособие написано на основе опыта преподавания автором курса теории вероятностей и математической статистики, рассчитанного на 48 часов аудиторных занятий в Институте радиоэлектроники и информационных технологий (ИРИТ-РТФ УрФУ). Этого времени хватает лишь на краткое введение в два огромных раздела математики, которые на современном уровне их развития используют очень тонкий и нетривиальный математический аппарат. Его освоение — непростая задача для студентов технических специальностей. При этом особую трудность представляет то, что с самого начала от них требуются особые мыслительные навыки — «вероятностная интуиция». Автор поставил перед собой задачу создать учебное пособие, которое помогло бы студенту познакомиться с нетривиальными понятиями изучаемой теории так, чтобы достаточно глубоко понять материал, по возможности избегая чрезмерного погружения в технические детали математического аппарата, и позволило бы преподавателю часть материала оставить студентам на самостоятельное изучение (такой материал входит в учебное пособие в виде приложений).

Как показывает опыт преподавания, 48-ми часовой курс распадается на три больших раздела. Первый представляет собой элементарное введение в основы теории вероятностей и обычно занимает 16 часов аудиторных занятий (8 лекционных и 8 практических). Ему соответствуют главы 1–4 учебного пособия. Задача этого раздела — научить студента основам математического моделирования случайных явлений, развить в нем «вероятностную интуицию». Второй раздел — самый большой. Он посвящен изучению более продвинутого математического аппарата, использующегося для анализа случайных величин и их распределений вероятностей. Обычно он занимает 24 часа аудиторных занятий (6 лекционных и 6 практических). Ему соответствуют главы 5–10. На последний раздел, посвященный основам математической статистики, остается лишь 8 часов аудиторных занятий. Ему соответствуют главы 11 и 12. Материал этого раздела удобно осваивать, сочетая изложение теории с проведением лабораторных работ в вычислительной среде *MatLab*. Приложение E содержит задания двух лабораторных работ по темам глав 11 и 12.

В конце каждой из первых девяти глав даны задачи для самостоятельного решения. В приложении F приведены ответы к задачам.

Глава 1. Классические модели теории вероятностей

1.1. Случайный эксперимент и случайные события

Теория вероятностей занимается построением и изучением математических моделей случайных явлений. При этом рассматриваются лишь те явления, которые вписываются в концепцию ***случайного эксперимента***.

Случайным экспериментом называют реальный или гипотетический (воображаемый) эксперимент, который может быть многократно повторен в одних и тех же условиях и может закончиться одним из некоторой заранее известной совокупности возможных исходов, при этом заранее неизвестно, каким именно.

В качестве примеров случайных экспериментов можно привести бросание монеты или игральной кости, раздачу игральных карт, соревнование по любому виду спорта, измерение какой-либо величины, при котором возникают случайные ошибки, стрельбу по мишени, торги на бирже и т. п.

В результате проведения случайного эксперимента могут происходить (или не происходить) различные события. В теории вероятностей события принято обозначать большими буквами латинского алфавита. Задавая события, мы будем описывать их словами, заключая описание в фигурные скобки.

■ Пример 1.1.1. Бросают игральную кость. В результате броска могут произойти (или не произойти) следующие события:

$A = \{\text{Выпала шестерка}\}$;

$B = \{\text{Выпало четное число}\}$;

$C = \{\text{Выпало меньше 4-х очков}\}$. ■

В основе построений теории вероятностей лежит эмпирически обнаруженное явление стабилизации частоты появления случайного события при увеличении числа повторений случайного эксперимента (***частотой наступления события*** при n повторениях случайного эксперимента называется величина $\frac{m}{n}$, где m — число наступлений данного события). Это явление дает основание для того, чтобы говорить об объективно существующей числовой характеристике события — его вероятности. С ростом числа повторений случайного эксперимента частота наступления события становится близкой к его вероятности. В этом смысле вероятность события характеризует его шансы произойти в данном случайном эксперименте.

1.2. Пространство элементарных событий

Базовым элементом математической модели случайного эксперимента является пространство элементарных событий.

Определение 1.2.1. ***Пространством элементарных событий*** для данного случайного эксперимента называют такой набор случайных событий, что в результате проведения этого эксперимента происходит одно и только одно событие из этого набора.

Пространство элементарных событий традиционно обозначают символом Ω. Его элементы обозначают ω и называют ***исходами эксперимента***.

■ Пример 1.2.1. Бросают монету. Этому случайному эксперименту можно поставить в соответствие пространство элементарных событий, состоящее из двух элементов:

$$\Omega = \{\omega_1 , \omega_2\} ,$$

где $\omega_1 = \{\text{Выпал герб}\}$, $\omega_2 = \{\text{Выпала решетка}\}$. ■

Ⓡ Необходимо отметить, что для построениия «разумной» математической модели, задавая пространство элементарных событий, важно исключить из рассмотрения те события, которые могут произойти на практике, но не представляют интереса и ведут к усложнению модели. Например, в результате броска монета может встать на ребро или укатиться, но включать эти исходы в пространство элементарных событий нецелесообразно.

■ Пример 1.2.2. Бросают игральную кость. Пространством элементарных исходов является набор

$$\Omega = \{\omega_1\,, \omega_2\,, \omega_3\,, \omega_4\,, \omega_5\,, \omega_6\}\,,$$

где $\omega_i = \{\text{выпало } i \text{ очков}\}\,, i = \overline{1, 6}$. ■

■ Пример 1.2.3. Бросают пару игральных костей. Для этого случайного эксперимента можно построить несколько различных пространств элементарных событий.

1. $\Omega_1 = \{\omega = \{x_1, x_2\}\,, \mid x_i \in \{1, 2, \ldots, 6\}\,, i = 1, 2\,, x_1 \le x_2\}$, где x_i — количество очков, выпавших на i-й кости. Таким образом, исходами считаем наборы из двух целых чисел, возможно, одинаковых (от 1 до 6).

2. $\Omega_2 = \{\omega = (x_1, x_2)\,, \mid x_i \in \{1, 2, \ldots, 6\}\,, i = 1, 2\}$, где x_i — количество очков, выпавших на i-й кости, т. е. исходы данного эксперимента мы отождествляем с упорядоченными парами чисел, различая, в отличие от Ω_1, например, исход $\omega = (1, 2)$ и исход $\omega = (2, 1)$.

3. $\Omega_3 = \{2\,, 3\,, \ldots, 12\}$, где $\omega \in \Omega$ — сумма очков, выпавших на игральных костях.

Все три набора исходов удовлетворяют определению пространства элементарных событий. Далее мы обсудим, какой из них и каким образом можно использовать для построения математической модели данного случайного эксперимента. ■

Ⓡ Дальше, так же, как в этом примере, используются круглые скобки для записи упорядоченных и фигурные скобки для записи неупорядоченных наборов элементов.

Элементы пространства элементарных событий — исходы — играют роль простейших атомарных событый в математической модели случайного эксперимента. Совокупности исходов образуют более сложные события.

Событиями называют подмножества (вообще говоря, не любые, но об этом будет сказано подробнее в следующей главе) пространства элементарных событий. При этом говорят, что событие $A \subseteq \Omega$ произошло, если случайный эксперимент закончился исходом $\omega \in \Omega$, принадлежащим A.

■ Пример 1.2.4. Бросают игральную кость. Пусть Ω определено как в примере 1.2.2. Тогда

$A = \{$выпало четное число$\} = \{\omega_2\,, \omega_4\,, \omega_6\}\,;$

$B = \{$выпало нечетное число$\} = \{\omega_1\,, \omega_3\,, \omega_5\}\,;$

$C = \{$выпало не больше трех очков$\} = \{\omega_1\,, \omega_2\,, \omega_3\}\,.$ ■

Множество Ω, рассматриваемое как свое собственное подмножество, также является событием. Его называют ***достоверным событием*** (оно происходит при любом исходе случайного эксперимента). Пустое множество $\varnothing$ называют ***невозможным событием***.

Определение 1.2.2. Если $A, B \subseteq \Omega$ — события и $A \subseteq B$, говорят, что событие A ***влечет за собой*** событие B или событие B является ***следствием события*** A.

Действительно, если $A \subseteq B$ и событие A произошло, т. е. случайный эксперимент закончился исходом $\omega \in A$, событие B также произошло.

1.3. Операции над событиями

Поскольку в теории вероятностей события представляют собой подмножества пространства элементарных исходов Ω, для них определены теоретико-множественные операции.

Определение 1.3.1. Пересечение событий $A \cap B$ называют ***произведением*** и обозначают AB, объединение событий $A \cup B$ называют ***суммой*** и обозначают $A + B$, теоретико-множественную разность $A \setminus B$ называют ***разностью*** событий и обозначают $A - B$, дополнение до множества A называют ***противоположным*** к A событием и обозначают $\overline{A}$ ($\overline{A} := \Omega - A$).

R В дальнейшем мы будем использовать как арифметические, так и теоретико-множественные обозначения для операций над событиями.

Из определения теоретико-множественных операций следует, что

- AB — событие, состоящее в том, что в результате случайного эксперимента произошли оба события (A и B);
- $A + B$ — событие, состоящее в том, что в результате случайного эксперимента произошло по крайней мере одно из событий (A или B);
- $A - B$ — событие, состоящее в том, что в результате случайного эксперимента произошло событие A и не произошло событие B;
- $\overline{A}$ — событие, состоящее в том, что в результате случайного эксперимента событие A не произошло.

Определение 1.3.2. События A и B называют ***несовместными***, если $AB = \varnothing$.

Математическая модель случайного эксперимента построена, если для всех подмножеств пространства элементарных исходов, являющихся событиями, определены вероятности. Вероятность события A обозначают $P(A)$. Это число из промежутка $[0; 1]$. Как отмечалось выше, оно характеризует шансы события A произойти в данном случайном эксперименте. Таким образом, P — функция множества. Рассмотрим как она определяется в двух классических математических моделях теории вероятностей — в схеме с конечным числом равновероятных исходов и в схеме с геометрическими вероятностями.

1.4. Схема с конечным числом равновероятных исходов

Пусть случайный эксперимент имеет конечное множество исходов $\Omega = \{\omega_1, \omega_1, \ldots, \omega_n\}$, при этом можно считать все эти исходы равновозможными, т. е. имеющими одинаковые шансы произойти (обычно это связано с такими свойствами рассматриваемых в модели объектов, как симметрия, однородность и т. п.). В такой ситуации естественно каждому исходу приписать одну и ту же вероятность, равную $\frac{1}{n}$ (общая «масса» вероятности, равная 1, делится поровну между всеми исходами). В результате для произвольного события $A = \{\omega_{k_1}, \ldots .\omega_{k_m}\} \subseteq \Omega$ вероятность оказывается равна $\frac{m}{n}$. Иначе говоря, вероятность определена на подмножествах Ω формулой

$$\boxed{P(A) = \frac{N(A)}{N(\Omega)}, \qquad (1.1)}$$

где $N(A)$ — количество элементов множества A (их называют исходами, ***благоприятными*** для A), $N(\Omega)$ — количество элементов пространства элементарных событий.

R Условие равновозможности исходов является крайне важным для корректного использования данной модели и формулы (1.1). Рассмотрим случайный эксперимент с бросанием пары игральных костей, для которого в примере 1.2.3 предложено три набора исходов, удовлетворяющих определению пространства элементарных событий. Из них равновероятными можно считать только элементы Ω_2, представляющие собой упорядоченные пары (x_1, x_2), где x_1 — число очков на 1-й кости, а x_2 — число очков на 2-й кости. То, что элементы пространства Ω_3 нельзя считать равновозможными исходами — очевидно. Например, сумму очков 7 можно получить разными способами, а 12 — только одним. Менее очевидно, что исходы пространства Ω_1 нельзя считать равновозможными, ведь, на первый взгляд, если игральные кости неразличимы, мы можем увидеть только какие числа выпали, а порядок при этом неважен. Тем не менее, элементы Ω_1 — неупорядоченные пары чисел — нельзя считать равновозможными исходами, так как две игральные кости это физически разные объекты (мы можем умозрительно раскрасить их, назвав одну черной, а другую — белой), а значит, к примеру, $(2, 1)$ и $(1, 2)$ — разные исходы, поэтому появление единицы и двойки имеет больше шансов, чем, например появление двух единиц.

R При вычислении вероятностей по формуле (1.1) для подсчета количества элементов того или иного множества используют приемы и формулы комбинаторики. Приложение А содержит краткое изложение некоторых базовых определений и формул этой теории, которые могут пригодиться при решении задач.

Задача 1.4.1. На шахматную доску наудачу поставили две ладьи. Какова вероятность того, что они смогут срубить друг друга?

Решение. Случайный эксперимент состоит в выборе наудачу двух полей на шахматной доске для постановки ладей. Предварительно перенумеровав поля (неважно каким образом), введем пространство элементрарных событий

$$\Omega = \Big\{\omega = (x, y) \,|\, x, y \in \{1, 2, \dots, 64\}, x \neq y\Big\},$$

где x — позиция первой ладьи, y — позиция второй ладьи (можно считать, что их ставят на доску одну за другой). Таким образом, исходы данного случайного эксперимента представляют собой размещения из 64 по 2 (см. Приложение A, Определение A.0.1), поэтому $N(\Omega) = A_{64}^2 = 64 \cdot 63$. Пусть A — событие, вероятность которого требуется найти:

$$A = \{\text{Ладьи могут срубить друг друга}\}.$$

Для подсчета исходов эксперимента, благоприятных для A, используем правило произведения. Для построения исхода, при котором ладьи могут срубить друг друга, первую из них можно поставить на любое из 64 полей доски, при этом вторую нужно поставить в ту же вертикаль или горизонталь, что можно сделать 14-ю способами. Таким образом, $N(A) = 64 \cdot 14$. По формуле (1.1) получаем $P(A) = \dfrac{14}{63} = \dfrac{2}{9}$. ■

Задача 1.4.2. Из урны, в которой содержатся n белых, m черных и k красных шаров, наудачу извлекают три шара. Какова вероятность того, что все они разного цвета? Какова вероятность того, что среди извлеченных есть шары только одного цвета, если $n \geq 3$, $m \geq 3$ и $k \geq 3$?

Решение. Исходом случайного эксперимента, состоящего в извлечении трех шаров наудачу из совокупности в $n+m+k$ шаров, является сочетание из $n+m+k$ шаров по три. Пространством элементарных событий Ω для данного случайного эксперимента будем считать совокупность всех таких сочетаний. Таким образом, $N(\Omega) = C^3_{n+m+k}$. Введем обозначения для событий:

$$A = \{\text{все извлеченные шары разных цветов}\},$$
$$B = \{\text{все извлеченные шары одного цвета}\}.$$

Чтобы построить сочетание, дающее исход, который принадлежит событию A, нам нужно выбрать по одному шару каждого цвета. При этом число способов выбрать белый шар равно n, черный — m и красный — k. По правилу произведения получаем $N(A) = mnk$. В результате

$$P(A) = \frac{N(A)}{N(\Omega)} = \frac{nmk}{C^3_{n+m+k}} = \\ = \frac{6nmk}{(n+m+k)(n+m+k-1)(n+m+k-2)}.$$

Для того чтобы построить сочетание, дающее исход, который принадлежит событию B, нужно выбрать либо только белые шары, что можно сделать C^3_n способами, так как каждый набор из трех белых шаров представляет собой сочетание из n имеющихся белых шаров по три, либо только черные шары, что можно сделать C^3_m способами, либо только красные шары, что можно сделать C^3_k способами. Применяя правило суммы, получаем $N(B) = C^3_n + C^3_m + C^3_k$. В результате

$$P(B) = \frac{N(B)}{N(\Omega)} = \frac{C^3_n + C^3_m + C^3_k}{C^3_{n+m+k}} = \\ = \frac{n(n-1)(n-2) + m(m-1)(m-2) + k(k-1)(k-2)}{(n+m+k)(n+m+k-1)(n+m+k-2)}. \quad \blacksquare$$

1.5. Схема с геометрическими вероятностями

Пусть исходы случайного эксперимента можно отождествить с точками некоторого измеримого множества $\Omega \subset \mathbb{R}^d$ так, что все исходы можно считать равновозможными. В этих условиях формула (1.1) неприменима, так как Ω является бесконечным множеством. Отвлекаясь от природы моделируемого случайного эксперимента, можно представлять себе его как выбор наудачу точки из Ω или бросание наудачу точки в Ω. Событие $A \subset \Omega$ происходит, если точка попадает в A. В такой ситуации естественно приписать A вероятность, пропорциональную мере этого множества. Таким образом, мы приходим к формуле

$$\boxed{P(A) = \frac{\mu(A)}{\mu(\Omega)}\,. \qquad (1.2)}$$

Задача 1.5.1. (задача о встрече)

Двое договорились о встрече в определенном месте в течение определенного часа. Какова вероятность того, что встреча произойдет, если каждый из них независимо от другого может прийти в условленное место в любой момент этого часа и будет ждать другого в течение 15 минут?

Решение. Случайный эксперимент в данном случае состоит в наблюдении за появлением двух человек в условленном месте в течение данного часа. Его исход можно описать, например, парой чисел x и y — моментами появления первого и второго участников эксперимента (измеряем их в часах от начала условленного часа). Таким образом, можно определить пространство элементарных событий положив

$$\Omega = \{\omega = (x, y) \,|\, x, y \in [0; 1]\}\,.$$

Поскольку все исходы равновозможны, можно отождествить данный случайный эксперимент с извлечением наудачу точки из квад-

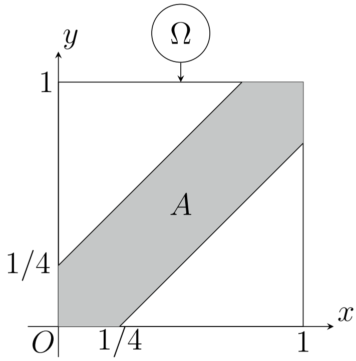


Рис. 1.1. Задача о встрече

рата Ω на координатной плоскости.

Пусть A — событие, о котором идет речь в условии задачи. Тогда

$$A = \{\text{Встреча произошла}\} = \Big\{\omega = (x, y) \in \Omega \,|\, |x - y| < 1/4\Big\}\,.$$

Таким образом, событие A — подмножество квадрата Ω, изображенное на рис. 1.1. Площадь области A равна $\dfrac{7}{16}$. Следовательно, получаем

$$P(A) = \frac{\mu(A)}{\mu(\Omega)} = \frac{7}{16}\,.$$

■

Ⓡ В схеме с геометрическими вероятностями, в отличие от схемы с конечным числом равновероятных исходов, вероятность определена фомулой (1.2) только для измеримых подмножеств Ω. Таким образом, множество событий, для которых определена вероятность, отличается от совокупности всех подмножеств пространства элементарных событий.

Ⓡ Исходы в схеме с геометрическими вероятностями равновозможны в том смысле, что все они равноправны с точки зрения возможности их получения в моделируемом эксперименте. При этом вероятность каждого исхода по формуле (1.2) равна нулю. Таким образом, в этой модели мы впервые сталкиваемся с событиями, которые не являются невозможными (отличными от $\varnothing$), но имеют нулевую вероятность.

Такие события принято называть ***практически невозможными***. Еще один пример практически невозможного события в схеме с геометрическими вероятностями:

$$B = \{\omega = (x, y) \in \Omega \,|\, x = y\}\,.$$

Его можно описать как событие, состоящее в том, что участники случайного эксперимента пришли в условленное место одновременно. На рис. 1.1 ему соответствует диагональ квадрата Ω, соединяющая точки $(0;0)$ и $(1;1)$, площадь которой равна нулю. Поэтому $P(B) = 0$. Такой результат может показаться странным с точки зрения практического опыта, однако на практике при проведении данного случайного эксперимента значения x и y оказываются не произвольными элементами отрезка $[0;1]$, множества мощности континуум, а элементами некоторого конечного дискретного подмножества этого отрезка, устройство которого определяется ценой деления шкалы прибора, с помощью которого фиксируется время. В результате одновременное прибытие участников это всегда событие вида

$$C = \{\omega = (x, y) \in \Omega \,|\, |x - y| < \varepsilon\}\,,$$

а значит имеет хоть и очень маленькую, но ненулевую вероятность:

$$P(C) = 1 - (1-\varepsilon)^2\,.$$

Задача 1.5.2. (задача Бюффона)

Бесконечная плоскость расчерчена параллельными прямыми, отстоящими друг от друга на расстояние L. На плоскость бросают иглу длины l $(l < L)$. Найти вероятность того, что игла пересечет одну из прямых.

Решение. Пусть r — расстояние от нижнего конца иглы до ближайшей к ней сверху прямой, а φ — угол между иглой и этой прямой, отсчитываемый против часовой стрелки (см. рис. 1.2). Исход случайного эксперимента, состоящего в бросании иглы на плоскость, однозначно определяется парой чисел (φ, r), где $\varphi \in [0;\pi)$, $r \in [0;L)$, поэтому положим $\Omega = [0;\pi) \times [0;L)$. Игла пересекает одну из параллельных прямых тогда и только тогда, когда выполнено условие $r \leq l\sin\varphi$, значит

$$\begin{aligned} A &= \{\text{игла пересекает одну из прямых}\} = \\ &= \Big\{\omega = (\varphi, r) \in \Omega \,|\, r \leq l\sin\varphi\Big\}\,. \end{aligned}$$

Событие A изображено на рис. 1.2. По формуле (1.2) находим

$$P(A) = \frac{\mu(A)}{\mu(\Omega)} = \frac{1}{\pi L} \int_0^{\pi} l \sin\varphi \, d\varphi = \frac{2l}{\pi L}\,. \qquad \blacksquare$$

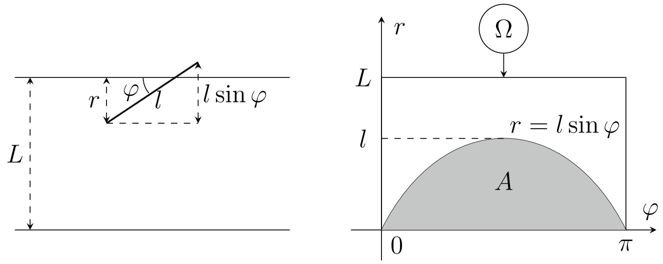


Рис. 1.2. Задача Бюффона

1.6. Задачи для самостоятельного решения

1. На десяти карточках написаны буквы слова «СТАТИСТИКА». Из этих 10 карточек наудачу извлекли 5 карточек. Найти вероятность того, что из отобранных карточек можно составить слово «ТАКСИ».
2. В урне находятся N шаров, из которых M — черные. Шары извлекаются друг за другом наудачу. Найти вероятность того, что k-й извлеченный шар — черный ($k = 1, 2, \ldots, N$).
3. За круглый стол в случайном порядке рассаживаются n женщин и n мужчин. Найти вероятность того, что рядом с каждой женщиной будут сидеть только мужчины, а рядом с каждым мужчиной — только женщины.
4. Найти вероятность того, что два ферзя, наудачу поставленные на шахматную доску, не смогут срубить друг друга.
5. Из 28 костей домино наудачу выбираются две. Найти вероятность того, что из них можно составить цепочку согласно правилам игры.
6. Числа p и q наудачу выбираются из отрезка $[-b; b]$. Найти вероятность того, что уравнение $x^2 + px + q = 0$ имеет действительные

корни. Найти предел, к которому стремится эта вероятность при $b \to \infty$.

7. На отрезке наудачу выбирают две точки, которые делят его на три части. Найти вероятность того, что из полученных отрезков можно построить треугольник.
8. Точку наудачу бросают в круг радиуса R с центром в начале координат. Найти вероятность того, что

   a) круг радиуса $r < R$ с центром в этой точке целиком содержится в большом круге;

   b) квадрат со стороной $r < R\sqrt{2}$ с центром в этой точке и сторонами, параллельными координатным осям, целиком содержится в круге.
9. Круг радиуса R, лежащий в плоскости, вращается в этой плоскости вокруг своего центра с постоянной угловой скоростью. Отрезок длиной $2h$ располагается в той же плоскости так, что прямая, проходящая через середину отрезка и центр круга, перпендикулярна отрезку, а расстояние от центра круга до отрезка равно l $(R < l)$. В случайный момент времени с окружности, ограничивающей круг, по касательной к ней слетает частица. Найти вероятность попадания этой частицы на отрезок.

Глава 2. Аксиоматика и формулы теории вероятностей

2.1. Вероятностное пространство

Из примеров, рассмотренных на предыдущей лекции, можно увидеть, что математическая модель случайного эксперимента построена тогда, когда описано Ω — пространство элементарных событий или исходов эксперимента — и на некоторой совокупности подмножеств множества Ω определена вероятность. При этом, как было видно в модели с геометрическими вероятностями, совокупность событий может отличаться от совокупности всех подмножеств множества Ω. Естественным требованием к ней является требование замкнутости относительно операций над событиями, т. е. относительно теоретико-множественных операций, а главное свойство, которым должна обладать вероятность — свойство аддитивности: $P(A+B) = P(A) + P(B)$ для несовместных событий. Таким образом, вероятность — мера, определенная на совокупности событий.

Все это формализуется с помощью аксиом, заложенных в определение центрального понятия теории вероятностей — понятия ***вероятностного пространства***.

Определение 2.1.1. Вероятностным пространством называется тройка $(\Omega, \mathcal{F}, P)$, где
Ω — ***пространство элементарных событий*** (см. Определе-

ние 1.2.1);

$\mathcal{F}$ — ***сигма-алгебра событий***, т. е. совокупность подмножеств множества Ω, удовлетворяющая аксиомам

$(\sigma - I)$ $\Omega \in \mathcal{F}$,

$(\sigma - II)$ $A, B \in \mathcal{F} \;\Rightarrow\; A \setminus B \in \mathcal{F}$,

$(\sigma - III)$ $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F} \;\Rightarrow\; \bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{F}$;

P — ***вероятностная мера***, т. е. функция $P : \mathcal{F} \to \mathbb{R}$, удовлетворяющая аксиомам

$(P - I)$ $\forall\, A \in \mathcal{F}\;\; P(A) \geq 0\,$,

$(P - II)$ $\forall\, \{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$

$$\left(A_i \cap A_j = \varnothing \text{ при } i \neq j \;\Rightarrow\; P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)\right),$$

$(P - III)$ $P(\Omega) = 1\,$.

Таким образом, вероятностное пространство представляет собой математическую модель случайного эксперимента.

R Аксиомы сигма-алгебры $(\sigma - I)$, $(\sigma - II)$ и $(\sigma - III)$ влекут за собой замкнутость совокупности $\mathcal{F}$ всех событий относительно операций сложения, умножения, вычитания и взятия противоположного события над любыми конечными и счетными совокупностями событий. Это следует из Предложения 2.1.1.

Предложение 2.1.1. Пусть $\mathcal{F}$ — сигма-алгебра подмножеств множества Ω. Тогда справедливы следующие утверждения:

a) $\varnothing \in \Omega$;

b) $A \in \mathcal{F} \;\Rightarrow\; \overline{A} \in \mathcal{F}$;

c) $\{A_i\}_{i=1}^{n} \subset \mathcal{F} \;\Rightarrow\; \bigcup\limits_{i=1}^{n} A_i \in \mathcal{F}$;

d) $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F} \;\Rightarrow\; \bigcap\limits_{i=1}^{\infty} A_i \in \mathcal{F}$;

e) $\{A_i\}_{i=1}^{n} \subset \mathcal{F} \;\Rightarrow\; \bigcap\limits_{i=1}^{n} A_i \in \mathcal{F}$.

Доказательство.

a) Из аксиом $(\sigma - I)$ и $(\sigma - II)$ следует: $\varnothing = \Omega \setminus \Omega \in \mathcal{F}$.

b) Для $A \in \mathcal{F}$ из аксиом $(\sigma - I)$ и $(\sigma - II)$ следует: $\overline{A} = \Omega \setminus A \in \mathcal{F}$.

c) Пусть $A_i = \varnothing$ при $i = n+1, n+2, \ldots$. Тогда для $\{A_i\}_{i=1}^{n} \subset \mathcal{F}$ из утверждения (a) и аксиомы $(\sigma - III)$ следует $\bigcup\limits_{i=1}^{n} A_i = \bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{F}$.

Утверждения (d) и (e) вытекают из утверждения (c) и аксиомы $(\sigma - III)$ соответственно, в силу закона де Моргана $\bigcap\limits_{i} A_i = \overline{\bigcup\limits_{i} \overline{A_i}}$ и утверждения (b). ■

▪ Пример 2.1.1. Схеме с конечным числом равновероятных исходов соответствует вероятностное пространство $(\Omega, \mathcal{F}, P)$, где пространство элементарных исходов конечно: $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$, $n \in \mathbb{N}$, сигма-алгебра $\mathcal{F}$ представляет собой совокупность всех подмножеств множества Ω, а вероятностная мера задана на $\mathcal{F}$ формулой (1.1). ▪

Заложенное в определении вероятностного пространства требование того, чтобы совокупность всех событий $\mathcal{F}$ была сигма-алгеброй подмножеств множества Ω естественно, так как довольно часто возникает необходимость рассматривать операции над бесконечными наборами событий. Например, в эксперименте, состоящем в бросании монеты до первого выпадения герба, где $A_i = \{$при i-м броске выпал герб$\}$, можно говорить о событии $\bigcup\limits_{i=1}^{\infty} A_i$, которое состоит в том, что герб выпал (при некотором броске).

> R Функция $\mu : \mathcal{F} \to \mathbb{R}$, определенная на сигма-алгебре $\mathcal{F}$ подмножеств некоторого множества и удовлетворяющая аксиомам $(P - I)$ и $(P - II)$ (***аксиома счетной, или сигма-аддитивности***) называется ***мерой на сигма-алгебре***. Вероятностная мера — мера на сигма-алгебре, удовлетворяющая аксиоме $(P - III)$.

▪ Пример 2.1.2. Схеме с геометрическими вероятностями в общем случае соответствует вероятностное пространство $(\Omega, \mathcal{F}, P)$, где Ω — измеримое по Лебегу подмножество пространства $\mathbb{R}^n$, $n \in \mathbb{N}$; $\mathcal{F}$ — сигма-алгебра измеримых по Лебегу подмножеств множества Ω; P — вероятностная мера, определенная на $\mathcal{F}$ формулой

$$P(A) = \frac{\mu(A)}{\mu(\Omega)}, \quad A \in \mathcal{F}, \tag{2.1}$$

где μ — мера Лебега в $\mathbb{R}^n$. Для случайного эксперимента, рассмотренного в задаче о встрече (Задача 1.5.1) $\Omega = [0;1] \times [0;1]$, $n = 2$. ▪

R При решении задачи 1.5.1 в качестве меры μ в формуле (2.1) использовалась мера Жордана на плоскости, которая представляет собой обобщение понятия площади (изначально определенной для всевозможных многоугольников) на совокупность фигур, измеримых по Жордану. Однако совокупность измеримых по Жордану подмножеств квадрата Ω не является сигма-алгеброй. Мера Лебега на плоскости также представляет собой обобщение понятия площади, но, в отличие от меры Жордана, ее область определения — совокупность измеримых по Лебегу подмножеств Ω — является σ-алгеброй. Измеримые по Жордану множества измеримы по Лебегу и на измеримых по Жордану множествах меры Жордана и Лебега совпадают. При этом встретить в приложениях конкретное множество измеримое по Лебегу, но неизмеримое по Жордану очень сложно. Так что с практической точки зрения использование меры Лебега не приносит ничего нового по сравнению с мерой Жордана. Тем не менее, тот факт, что в результате применения любых теоретико-множественных операций к конечным и счетным наборам измеримых по Лебегу множеств получаются измеримые по Лебегу множества, играет фундаментальную роль при построении математического аппарата теории вероятностей. Подробное изложение теории меры и конструкции меры Лебега можно найти в книге [6].

Рассмотрим свойства вероятностной меры, вытекающие из аксиоматики вероятностного пространства, которые используются для решения всевозможных задач о «пересчете вероятностей». Это задачи, в которых по известным вероятностям некоторого набора событий, обычно достаточно простых, требуется найти вероятности других, более сложно устроенных событий. При этом предполагается, что вероятностное пространство, соответствующее рассматриваемому случайному эксперименту, существует, но, в отличие от задач, решаемых с использованием схемы с конечным числом равновероятных исходов или схемы с геометрическими вероятностями, явное описание этого вероятностного пространства не требуется.

2.2. Свойства вероятности

Предложение 2.2.1. Вероятностная мера обладает следующими свойствами:

(P-1) $P(\varnothing) = 0$;

(P-2) ***(Конечная аддитивность вероятности)***

$$\forall \{A_i\}_{i=1}^n \subset \mathcal{F} \; (i \neq j \Rightarrow A_i A_j = \varnothing) \Rightarrow P\left(\sum_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i)\,;$$

(P-3) $\forall\, A, B \in \mathcal{F} \;\; A \subseteq B \;\Rightarrow\; P(B - A) = P(B) - P(A)\,;$

(P-4) $\forall\, A, B \in \mathcal{F} \;\; A \subseteq B \;\Rightarrow\; P(A) \leq P(B)\,;$

(P-5) $\forall\, A \in \mathcal{F} \;\; 0 \leq P(A) \leq 1\,;$

(P-6) $\forall\, A \in \mathcal{F} \;\; P(\overline{A}) = 1 - P(A)\,;$

(P-7) ***(Формула сложения вероятностей)*** $\forall \{A_i\}_{i=1}^n \subset \mathcal{F}$

$$P\left(\sum_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i A_j) + \\ + \sum_{1 \leq i < j < k \leq n} P(A_i A_j A_k) - \dots + (-1)^{n-1} P\left(\prod_{i=1}^n A_i\right)\;;$$

(P-8) $\forall \{A_i\}_{i=1}^n \subset \mathcal{F} \quad P\left(\sum_{i=1}^n A_i\right) \leq \sum_{i=1}^n P(A_i)\,;$

(P-9) ***(Непрерывность вероятности)*** $\forall \{A_i\}_{i=1}^\infty \subset \mathcal{F}$:

1. $A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots \;\Rightarrow\; \lim\limits_{n\to\infty} P(A_n) = P\left(\sum_{i=1}^\infty A_i\right)\;;$

2. $A_1 \supseteq A_2 \supseteq A_3 \supseteq \dots \;\Rightarrow\; \lim\limits_{n\to\infty} P(A_n) = P\left(\prod_{i=1}^\infty A_i\right)\;.$

Доказательство. Для доказательства свойства (P-1) возьмем произвольный набор $\{A_i\}_{i=1}^\infty \subset \mathcal{F}$, состоящий из попарно несовместных событий. Положим $A_0 = \varnothing$. По аксиоме $(P - II)$ имеем:

$$P\left(\sum_{i=1}^\infty A_i\right) = P\left(\sum_{i=0}^\infty A_i\right) = \sum_{i=0}^\infty P(A_i) = \\ = P(\varnothing) + \sum_{i=1}^\infty P(A_i) = P(\varnothing) + P\left(\sum_{i=1}^\infty A_i\right)\;,$$

откуда следует $P(\varnothing) = 0$.

Докажем (Р-2). Пусть $\{A_i\}_{i=1}^n \subset \mathcal{F}$ $(i \neq j \;\Rightarrow\; A_iA_j = \varnothing)$. Положим $A_i = \varnothing$ при $i \in \mathbb{N}$ таких, что $i > n$. Тогда по аксиоме $(P - II)$ получим

$$P\left(\sum_{i=1}^{n} A_i\right) = P\left(\sum_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) = \sum_{i=1}^{n} P(A_i)\,,$$

так как $P(A_i) = P(\varnothing) = 0$ при $i > n$.

Пусть $A, B \in \mathcal{F}$, $A \subseteq B$. Тогда $B = A + (B - A)$, при этом события A и $B-A$ несовместны. По свойству конечной аддитивности вероятности (Р-2) $P(B) = P(A)+P(B-A)$. Таким образом, доказали (Р-3).

Свойство (Р-4) очевидным образом следует из (Р-3).

Свойство (Р-5) следует из (Р-4), (Р-1) и аксиомы $(P - III)$.

Свойство (Р-6) следует из равенства $A + \overline{A} = \Omega$, несовместности A и $\overline{A}$ и конечной аддитивности вероятности (Р-2).

Свойство (Р-7) доказывается по индукции. Для двух слагаемых, используя свойства (Р-2) и (Р-3), имеем:

$$\begin{aligned} P(A_1 + A_2) &= P\big((A_1 - A_1A_2) + (A_2 - A_1A_2) + A_1A_2\big) = \\ &= P(A_1 - A_1A_2) + P(A_2 - A_1A_2) + P(A_1A_2) = \\ &= P(A_1) - P(A_1A_2) + P(A_2) - P(A_1A_2) + P(A_1A_2) = \\ &= P(A_1) + P(A_2) - P(A_1A_2)\,. \end{aligned}$$

База индукции доказана. Предположим, что формула верна при некотором n. Докажем ее для $n + 1$ слагаемого. Используя базу индукции, получаем

$$\begin{aligned} P\left(\sum_{i=1}^{n+1} A_i\right) &= P\left(\sum_{i=1}^{n} A_i + A_{n+1}\right) = P\left(\sum_{i=1}^{n} A_i\right) + P(A_{n+1}) - \\ &- P\left(A_{n+1}\sum_{i=1}^{n} A_i\right) = P\left(\sum_{i=1}^{n} A_i\right) + P(A_{n+1}) - P\left(\sum_{i=1}^{n} A_iA_{n+1}\right). \end{aligned}$$

Преобразуя первое и последнее слагаемые с помощью предположения индукции, получаем

$$P\left(\sum_{i=1}^{n+1} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{1\le i<j\le n} P(A_iA_j) + \\ + \sum_{1\le i<j<k\le n} P(A_iA_jA_k) - \dots + (-1)^{n-1} P\left(\prod_{i=1}^{n} A_i\right) + \\ + P(A_{n+1}) - \sum_{i=1}^{n} P(A_iA_{n+1}) + \sum_{1\le i<j\le n} P(A_iA_jA_{n+1}) - \\ - \sum_{1\le i<j<k\le n} P(A_iA_jA_kA_{n+1}) + \dots + (-1)^{n} P\left(\prod_{i=1}^{n+1} A_i\right).$$

Группируя вероятности событий $A_1, A_2, \dots, A_{n+1}$, вероятности попарных произведений, произведений всевозможных троек этих событий и т. д., получаем окончательно

$$P\left(\sum_{i=1}^{n+1} A_i\right) = \sum_{i=1}^{n+1} P(A_i) - \sum_{1\le i<j\le n+1} P(A_iA_j) + \\ + \sum_{1\le i<j<k\le n+1} P(A_iA_jA_k) - \dots + (-1)^{n} P\left(\prod_{i=1}^{n+1} A_i\right).$$

Таким образом, доказан шаг индукции.

Свойство (P-8) доказывается также по индукции. База индукции:

$$P(A_1 + A_2) = P(A_1) + P(A_2) - P(A_1A_2) \le P(A_1) + P(A_2).$$

Предположим, свойство доказано для некоторого n. Тогда

$$P\left(\sum_{i=1}^{n+1} A_i\right) = P\left(\sum_{i=1}^{n} A_i\right) + P(A_{n+1}) - P\left(A_{n+1}\sum_{i=1}^{n} A_i\right) \le \sum_{i=1}^{n+1} P(A_i).$$

Здесь мы отбросили последнее слагаемое, так как оно неположительно и воспользовались предположением индукции.

Докажем свойство (P-9). Пусть $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$ и $A_n \subseteq A_{n+1}$ для любого $n \in \mathbb{N}$. Тогда, пользуясь аксиомой $(P{-}II)$ и тем, что события

$A_1, (A_2 - A_1), (A_3 - A_2), \ldots$ попарно несовместны, получаем:

$$P\left(\sum_{i=1}^{\infty} A_i\right) = P\left(A_1 + \sum_{i=2}^{\infty}(A_i - A_{i-1})\right) =$$
$$= P(A_1) + \sum_{i=2}^{\infty} P(A_i - A_{i-1}) =$$
$$= \lim_{n\to\infty}\left(P(A_1) + \sum_{i=2}^{n} P(A_i - A_{i-1})\right) =$$
$$= \lim_{n\to\infty} P\left(A_1 + \sum_{i=2}^{n}(A_i - A_{i-1})\right) = \lim_{n\to\infty} P(A_n)\,.$$

Пусть теперь $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$ и $A_n \supseteq A_{n+1}$ для любого $n \in \mathbb{N}$. Тогда, пользуясь законом де Моргана, уже доказанным предельным соотношением и тем, что $\overline{A}_n \subseteq \overline{A}_{n+1}$ для любого $n \in \mathbb{N}$, получаем:

$$P\left(\prod_{i=1}^{\infty} A_i\right) = 1 - P\left(\overline{\prod_{i=1}^{\infty} A_i}\right) = 1 - P\left(\sum_{i=1}^{\infty} \overline{A}_i\right) = 1 - \lim_{n\to\infty} P\big(\overline{A}_n\big) =$$
$$= \lim_{n\to\infty}\big(1 - P\big(\overline{A}_n\big)\big) = \lim_{n\to\infty} P(A_n)\,. \qquad \blacksquare$$

R Свойство конечной аддитивности вероятности (P-2) представлят собой частный случай формулы сложения (P-7). Действительно, если события набора $\{A_i\}_{i=1}^{n}$ попарно несовместны, их всевозможные произведения представляют собой невозможные события, вероятности которых равны нулю.

Задача 2.2.1. При входе в клуб джентльмены сдают шляпы в гардероб. Сработала пожарная сигнализация и, эвакуируясь, каждый из n членов клуба в спешке берет шляпу наудачу. Какова вероятность того, что, по крайней мере, один из них взял свою шляпу?

Решение. Пронумеруем членов клуба, присвоим те же номера их шляпам и рассмотрим события

$$A = \{\text{по крайней мере один член клуба взял свою шляпу}\}\,,$$

$$A_i = \{\ i\text{-й член клуба взял свою } (i\text{-ю}) \text{ шляпу}\}\,.$$

Справедливо равенство

$$A = \sum_{i=1}^{n} A_i \,.$$

Поскольку события A_i не являются попарно несовместными, воспользуемся формулой сложения (Р-7). Вероятности произведений событий вида $A_{i_1} A_{i_2} \ldots A_{i_p}$, где $1 \le i_1 < i_2 < \cdots < i_p \le n$ найдем, воспользовавшись схемой с конечным числом равновероятных исходов. Исход случайного экспеpемента, в результате которого члены клуба берут шляпы, выбирая их наудачу, можно отождествить с перестановкой из n элементов — номеров шляп. Таким образом, пространство элементарных событий для него имеет вид:

$$\Omega = \big\{ \omega = (i_1, i_2, \ldots, i_n) \,|\, i_k \in \{1, 2, \ldots, n\} \big\} \,.$$

Здесь i_k — номер шляпы, которую взял k-й джентльмен. В результате $N(\Omega) = n!$, а $N(A_{i_1} A_{i_2} \ldots A_{i_p}) = (n-p)!$, поскольку если члены клуба с номерами $i_1, i_2, \ldots, i_p$ взяли свои шляпы, то количество способов распределить остальные шляпы равно числу перестановок из $n - p$ элементов. В результате

$$P(A_{i_1} A_{i_2} \ldots A_{i_p}) = \frac{N(A_{i_1} A_{i_2} \ldots A_{i_p})}{N(\Omega)} = \frac{(n-p)!}{n!}$$

и эти вероятности не зависят от выбора номеров $i_1, i_2, \ldots, i_p$. Количество слагаемых в сумме вида

$$\sum_{1 \le i_1 < i_2 < \cdots < i_p \le n} P(A_{i_1} A_{i_2} \ldots A_{i_p})$$

равно количеству способов, которыми можно выбрать набор номеров $\{i_1, i_2, \ldots, i_p\}$ из совокупности $\{1, 2, \ldots, n\}$, т. е. C_n^p — количеству сочетаний из n элементов по p. Следовательно,

$$P(A) = P\left(\sum_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{1\le i<j\le n} P(A_iA_j) +$$
$$+ \sum_{1\le i<j<k\le n} P(A_iA_jA_k) - \dots + (-1)^{n-1}P\left(\prod_{i=1}^{n} A_i\right) =$$
$$= C_n^1\frac{(n-1)!}{n!} - C_n^2\frac{(n-2)!}{n!} + C_n^3\frac{(n-3)!}{n!} + \dots + (-1)^{n-1}\frac{1}{n!} =$$
$$= 1 - \frac{1}{2!} + \frac{1}{3!} - \dots + (-1)^{n-1}\frac{1}{n!}\,.$$

Заметим, что $\lim\limits_{n\to\infty}\left(1 - \frac{1}{2!} + \frac{1}{3!} - \dots + (-1)^{n-1}\frac{1}{n!}\right) = 1 - e^{-1}$, так что при больших n вероятность события A приближается к значению $1 - e^{-1} \approx 0,6321$, не зависящему от n. ■

Для того чтобы получить еще одну из базовых формул теории вероятностей, а именно формулу умножения, нам понадобятся понятия ***условной вероятности*** и ***независимости событий***.

2.3. Условная вероятность и независимость событий

Пусть $(\Omega, \mathcal{F}, P)$ — некоторое вероятностное пространство.

Определение 2.3.1. Пусть $A, B \in \mathcal{F}$, $P(B) \neq 0$. ***Условной вероятностью*** события A относительно события B называют величину

$$P(A|B) := \frac{P(AB)}{P(B)}\,. \tag{2.2}$$

(R) Величину $P(A|B)$ называют еще вероятностью события A при условии, что произошло событие B.

Мотивацией для такого определения могут служить следующие примеры.

1. Пусть вероятностное пространство $(\Omega, \mathcal{F}, P)$ соответствует схеме с конечным числом равновероятных исходов (см. Пример 2.1.1).

Предположим, известно, что событие B произошло, и нас интересует вероятность события A с учетом этой информации. Естественно в таком случае принять B за новое пространство элементарных событий. Поскольку A в этой ситуации наступает только в том случае, когда случайный эксперимент заканчивается одним из исходов, принадлежащих B, который одновременно принадлежит A, получаем

$$P(A|B) = \frac{N(AB)}{N(B)} = \frac{N(AB)/N(\Omega)}{N(B)/N(\Omega)} = \frac{P(AB)}{P(B)}\,.$$

2. Пусть $(\Omega, \mathcal{F}, P)$ — вероятностное пространство, соответствующее схеме с геометрическими вероятностями (см. Пример 2.1.2). Рассуждая так же, как в предыдущем примере, получаем

$$P(A|B) = \frac{\mu(AB)}{\mu(B)} = \frac{\mu(AB)/\mu(\Omega)}{\mu(B)/\mu(\Omega)} = \frac{P(AB)}{P(B)}\,.$$

R Заметим, что определение условной вероятности (Определение 2.3.1) работает, только если вероятность условия отлична от нуля: $P(B) \neq 0$. При этом имеет смысл говорить об условной вероятности и в ситуации, когда $P(B) = 0$. Например, пусть $(\Omega, \mathcal{F}, P)$ — вероятностное пространство, соответствующее схеме с геометрическими вероятностями, где $\Omega = [0;1] \times [0;1]$. Представим, что мы рассматриваем случайный эксперимент, состоящий в извлечении наудачу точки из квадрата Ω. Пусть l — некоторая кривая, лежащая в Ω, а $l_1 \subset l$ — некоторая дуга этой кривой. Рассмотрим события

$$B = \{\text{выбранная точка принадлежит } l\}\,,$$
$$A = \{\text{выбранная точка принадлежит } l_1\}\,.$$

Тогда $P(B) = 0$, а значит условная вероятность $P(A|B)$ по формуле (2.2) не определена, хотя она имеет практический смысл в контексте данного случайного эксперимента. В лекции 8 будет рассмотрен математический аппарат, с помощью которого условная вероятность может быть определена и относительно условия, вероятность которого равна нулю.

Из формулы (2.2) следует формула

$$P(AB) = P(B)P(A|B)\,. \tag{2.3}$$

Ее можно обобщить на любое конечное множество событий.

Предложение 2.3.1. (Формула умножения вероятностей).

$$\forall\ \{A_i\}_{i=1}^n \subset \mathcal{F} \quad P(A_1A_2\dots A_n) = \\ = P(A_1)P(A_2|A_1)P(A_3|A_2A_1)\dots P(A_n|A_1\dots A_{n-1})\,. \tag{2.4}$$

Доказательство. Последовательно применяя формулу (2.3), получаем:

$$P(A_1A_2\dots A_n) = P((A_1A_2\dots A_{n-1})\cdot A_n) = \\ = P((A_1A_2\dots A_{n-2})\cdot A_{n-1})P(A_n|A_1\dots A_{n-1}) = \\ = P((A_1A_2\dots A_{n-3})\cdot A_{n-2})P(A_{n-1}|A_1\dots A_{n-2})P(A_n|A_1\dots A_{n-1}) = \\ = \dots = P(A_1)P(A_2|A_1)P(A_3|A_2A_1)\dots P(A_n|A_1\dots A_{n-1})\,. \qquad \blacksquare$$

Определение 2.3.2. События A и B называются ***независимыми***, если

$$P(AB) = P(A)P(B)\,.$$

Из формулы (2.2) очевидным образом вытекает следующее предложение, проясняющее смысл независимости.

Предложение 2.3.2. Пусть $A, B \in \mathcal{F}$, $P(A) \neq 0$ $P(B) \neq 0$, тогда независимость событий A и B эквивалентна каждому из следующих условий:

1. $P(A) = P(A|B)$;
2. $P(B) = P(B|A)\,.$

Таким образом, независимость событий означает, что наступление одного из них не меняет шансы наступить или не наступить для другого: условные вероятности $P(A|B)$ и $P(B|A)$ совпадают с безусловными — $P(A)$ и $P(B)$ соответственно.

На наборы, состоящие из более двух событий, понятие независимости распространяется следующим образом.

Определение 2.3.3. События $A_1, A_2, \dots, A_n \subset \mathcal{F}$ называются ***независимыми***, если для любого набора индексов i_k, таких, что $1 \leq i_1 <$

$i_2 < \ldots, i_p \le n$, выполняется равенство

$$P(A_{i_1}A_{i_2}\ldots A_{i_p}) = P(A_{i_1})P(A_{i_2})\ldots P(A_{i_p}). \tag{2.5}$$

R Заметим, что независимость набора из $n > 2$ событий в смысле определения 2.3.3, которую часто называют ***независимость в совокупности***, влечет за собой, в частности, попарную независимость этих событий в смысле определения 2.3.2. При этом независимость в совокупности является более сильным свойством, чем ***попарная независимость***. Это показывает следующий пример.

■ Пример 2.3.1. (Пирамида Бернштейна). Представим себе правильную треугольную пирамиду, три грани которой раскрашены в разные цвета: красный, синий и зеленый соответственно, а окраска четвертой содержит все три цвета. Пирамиду бросают наудачу, в результате чего она с равными вероятностями падает на одну из четырех граней. Рассмотрим события:

$$A = \{\text{Пирамида упала на грань, содержащую красный цвет}\},$$
$$B = \{\text{Пирамида упала на грань, содержащую синий цвет}\},$$
$$C = \{\text{Пирамида упала на грань, содержащую зеленый цвет}\}.$$

Имеем:

$$P(A) = P(B) = P(C) = \frac{1}{2}.$$

При этом

$$P(AB) = P(AC) = P(BC) = P(ABC) = \frac{1}{4}.$$

Таким образом, выполнены равенства

$$P(AB) = P(A)P(B),\ P(AC) = P(A)P(C),\ P(BC) = P(B)P(C),$$

а значит, события A, B и C являются попарно независимыми. При этом

$$P(ABC) = \frac{1}{4} \ne P(A)P(B)P(C) = \frac{1}{8},$$

поэтому эти события не являются независимыми в совокупности. ■

2.4. Задачи для самостоятельного решения

1. Для того чтобы разрушить мост, на него независимо друг от друга сбросили три бомбы с вероятностями попадания $0,1; 0,3$ и $0,4$.

   a) Какова вероятность того, что мост разрушен, если для этого нужно попадание по крайней мере одной бомбы?

   b) Какова вероятность того, что мост будет разрушен, если для этого нужно попадание по крайней мере двух бомб?
2. Из урны, содержащей 3 белых и 2 черных шара, по схеме случайного выбора без возвращения последовательно извлекаются шары. Найти вероятность p_k того, что черный шар впервые появится при k-м извлечении ($k = 1, 2, 3, 4$).
3. Собираясь в поход, студент положил в каждый из 20 карманов своей куртки по прянику. Каждый раз, собираясь подкрепиться, он начинает в случайном порядке просматривать карманы до тех пор, пока не найдет очередной пряник. Найти

   a) вероятность того, что поиск k-го пряника начнется с пустого кармана;

   b) вероятность того, что первые k пряников будут найдены с первой попытки.
4. Игрок должен сыграть матч, состоящий из трех партий, с двумя игроками (сильным и слабым), играя с ними по очереди. Матч считается выигранным, если он одержал две победы подряд. С кем он должен сыграть первую партию, чтобы вероятность выиграть матч для него была максимальной?
5. Из множества чисел $\{1, 2, 3, \ldots, N\}$ по схеме случайного выбора без возвращения выбираются три числа. Найти условную вероятность того, что третье число попадает в интервал, образованный первыми двумя, если известно, что первое число меньше второго.
6. Из 100 карточек с числами 00, 01, ... , 98, 99 случайно выбирается одна. Пусть η_1 — сумма, а η_2 — произведение цифр на выбранной карточке. Найти условную вероятность $P(\eta_1 = i \,|\, \eta_2 = 0\}$, $i = 0, 1, 2, \ldots, 18$.
7. Докажите, что если события A и B независимы, то независимы-

ми являются события

a) $\overline{A}$ и B,

b) $\overline{A}$ и $\overline{B}$.

8. Случайная точка $\omega = (x_1, x_2)$ наудачу выбирается из квадрата $[0; 1] \times [0; 1]$. Пусть $A_1 = \{x_1 \leq 1/2\}$, $A_2 = \{x_2 \leq 1/2\}$, $A_3 = \{(x_1 - 1/2)(x_2 - 1/2) < 0\}$.

   a) Являются ли события A_1, A_2, A_3 попарно независимыми?

   b) Являются ли они независимыми в совокупности?

9. Игроки поочередно бросают монету. Выигрывает тот, у кого раньше появится герб.

   a) Найти вероятность выигрыша для каждого из игроков, если играют двое.

   b) Найти вероятность выигрыша для каждого из игроков, если играют трое.

Глава 3. Формула полной вероятности и формула Байеса

В предыдущей главе мы рассмотрели аксиоматику теории вероятностей, понятия условной вероятности и независимости событий и получили основные формулы, главными из которых являются формула сложения вероятностей (Р-7) и формула умножения вероятностей (2.3). Каждая из них принимает более простой вид, если рассматриваемые события удовлетворяют некоторым условиям, а именно формула сложения вероятностей превращается в формулу (Р-2), если события попарно несовместны, а формула умножения вероятностей превращается в формулу (2.5), если события независимы. В этой главе мы рассмотрим еще две важных формулы, которые часто работают при пересчете вероятностей.

3.1. Формула полной вероятности

Очень часто возникает необходимость вычисления вероятности события, наступление которого связано с рядом обстоятельств. Если известны вероятности их наступления и условные вероятности интересующего нас события относительно этих обстоятельств, можно использовать так называемую ***формулу полной вероятности***. Для ее формулировки и вывода нам потребуется следующее важное понятие.

Определение 3.1.1. Набор событий $\{B_i\}_{i\in\mathcal{I}} \subset \mathcal{F}$, где $\mathcal{I}$ — набор индексов ($\mathcal{I} = \{1, 2, \dots, n\}$, $n \in \mathbb{N}$, или $\mathcal{I} = \mathbb{N}$) называется ***полной группой событий***, если

$$\bigsqcup_{i\in\mathcal{I}} B_i = \Omega\,.$$

Здесь используется символ $\bigsqcup$, которым обозначают ***дизъюнктное объединение***, т. е. объединение попарно непересекающихся множеств.

Таким образом, набор событий $\{B_i\}_{i\in\mathcal{I}} \subset \mathcal{F}$ образует полную группу, если он удовлетворяет двум требованиям:

- события набора попарно несовместны : $\quad B_iB_k = \varnothing$ при $i \neq k$;
- $\sum\limits_{i\in\mathcal{I}} B_i = \Omega\,.$

Это означает, что при проведении случайного эксперимента происходит одно и только одно из этих событий.

R Полная группа событий может быть как конечной, так и счетной.

Предложение 3.1.1. (Формула полной вероятности). Пусть $\mathcal{I}$ — конечный или счетный набор индексов, $\{B_i\}_{i\in\mathcal{I}} \subset \mathcal{F}$ — полная группа событий. Тогда

$$\forall\ A \in \mathcal{F} \quad P(A) = \sum_{i\in\mathcal{I}} P(B_k)P(A|B_k)\,. \tag{3.1}$$

Доказательство. Пользуясь свойствами полной группы событий и законом де Моргана $\big(A \bigcap (\bigcup_i B_i) = \bigcup_i (A \bigcap B_i)\big)$, получаем разложение события A в сумму попарно несовместных событий:

$$A = A\Omega = A\sum_{i\in\mathcal{I}} B_i = \sum_{i\in\mathcal{I}} AB_i\,.$$

Далее, пользуясь аддитивностью вероятности (счетной или конечной) и формулой умножения вероятностей (2.3), получаем

$$P(A) = P\left(\sum_{i\in\mathcal{I}} AB_i\right) = \sum_{i\in\mathcal{I}} P\left(AB_i\right) = \sum_{i\in\mathcal{I}} P(B_k)P(A|B_k)\,. \qquad \blacksquare$$

Задача 3.1.1. Есть пять обычных игральных костей и две, на гранях которых две единицы, две двойки и две тройки. Наудачу выбирают одну из игральных костей и бросают. Какова вероятность того, что выпадет четное число?

Решение. Пусть $A = \{\text{выпало четное число}\}$. Шансы получить четное число зависят от того, какая кость извлечена на первом шаге случайного эксперимента. Рассмотрим события

$$B_1 = \{\text{выбрана обычная игральная кость}\},$$
$$B_2 = \{\text{выбрана игральная кость, содержащая 1, 2 и 3}\}.$$

Они образуют полную группу событий. По формуле полной вероятности получаем:

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) = \frac{5}{7}\cdot\frac{1}{2} + \frac{2}{7}\cdot\frac{1}{3} = \frac{19}{42}. \quad \blacksquare$$

Задача 3.1.2. Двое поочередно бросают монету. Выигрывает тот, у кого первого выпадет герб. Найти вероятности выигрыша для каждого из игроков.

Решение. Рассмотрим способ решения этой задачи (см. задачу 9 (а) из раздела 2.4), основанный на формуле полной вероятности. Пусть $A = \{\text{выиграл 1-й игрок}\}$. Обозначим через x вероятность этого события. Рассмотрим события

$$B_1 = \{\text{При первом броске выпал герб}\},$$
$$B_2 = \{\text{При первом броске выпала решетка, а при втором — герб}\},$$
$$B_3 = \{\text{При первом и втором бросках выпала решетка}\}.$$

Они образуют полную группу событий. Для условных вероятностей выполняются равенства:

$$P(A|B_1) = 1, \quad P(A|B_2) = 0, \quad P(A|B_3) = x.$$

Последнее из них верно в силу того, что если произошло событие B_3, т. е., начав игру, первый игрок получил решетку, а затем второй игрок бросил монету и тоже получил решетку, право броска переходит к первому игроку, при этом его шансы победить такие же, как в начале игры. В результате, по формуле полной вероятности, мы получаем уравнение

$$x = P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3) = \\ = \frac{1}{2} + \frac{x}{4}\,.$$

Отсюда находим $x = \dfrac{2}{3}$, а вероятность выигрыша второго игрока равна $\dfrac{1}{3}$. ■

3.2. Формула Байеса

Предположим, что полная группа событий $\{B_i\}_{i\in\mathcal{I}}$ представляет собой набор гипотез о результате некоторого случайного эксперимента. До его проведения у нас есть некоторая информация о нем — мы знаем вероятности этих гипотез (вероятности $P(B_i)$ при этом называют «априорными»). Пусть теперь эксперимент проведен, мы не знаем, какая из гипотез верна, но при этом известно, что произошло некоторое событие A. Это можно интерпретировать как новую информацию о данном случайном эксперименте и задаться вопросом: как оценить вероятности гипотез B_i с учетом этой информации? Иначе говоря, чему равны условные вероятности $P(B_i|A)$?

Поставленную таким образом задачу ***о корректировке априорных вероятностей с учетом наблюдаемого события*** позволяет решить формула Байеса.

Предложение 3.2.1. (Формула Байеса). Пусть $\mathcal{I}$ — конечный или счетный набор индексов, $\{B_i\}_{i\in\mathcal{I}} \subset \mathcal{F}$ — полная группа событий. Тогда

$$\forall\ A \in \mathcal{F}\,, k \in \mathcal{I} \quad P(B_k|A) = \frac{P(B_k)P(A|B_k)}{\sum\limits_{i\in\mathcal{I}} P(B_k)P(A|B_k)}\,. \tag{3.2}$$

Доказательство. По определению условной вероятности (определение 2.3.1)

$$P(B_k|A) = \frac{P(B_kA)}{P(A)}\,.$$

Преобразуя числитель по формуле умножения вероятностей (2.3), а знаменатель — по формуле полной вероятности (3.1), получаем правую часть формулы (3.2). ■

R Используемая в доказательстве формула (2.2), а следовательно, и формула Байеса работает только при условии $P(A) \neq 0$.

Задача 3.2.1. По шоссе, на котором находится бензозаправка, едет поток машин, в котором число грузовых машин относится к числу легковых как 3 : 2. Известно, что на заправку заезжает в среднем одна из десяти проезжающих мимо грузовых машин и одна из пяти легковых. Найти вероятность того, что машина, заехавшая на заправку, — грузовик.

Решение. Рассматриваемый случайный эксперимент состоит в наблюдении за проезжающей мимо бензозаправки машиной (каждая новая машина — очередной эксперимент). Наблюдаемое событие:

$$A = \{\text{машина заехала на заправку}\}\,.$$

Имеется две гипотезы:

$$B_1 = \{\text{машина грузовая}\}\,,\quad B_2 = \{\text{машина легковая}\}\,.$$

Они образуют полную группу событий. Их вероятности:

$$P(B_1) = \frac{3}{5}\,,\quad P(B_2) = \frac{2}{5}\,.$$

По формуле Байеса находим искомую вероятность:

$$P(B_1|A) = \frac{P(B_1)P(A|B_1)}{P(B_1)P(A|B_1)+P(B_2)P(A|B_2)} = \frac{\frac{3}{5}\cdot\frac{1}{10}}{\frac{3}{5}\cdot\frac{1}{10}+\frac{2}{5}\cdot\frac{1}{5}} = \frac{3}{7}.$$ ■

3.3. Задачи для самостоятельного решения

1. В одной урне находятся 1 белый и 9 черных шаров, а в другой — 1 черный и 5 белых шаров. Из каждой урны взяли по одному шару, выбрав их наудачу, а оставшиеся положили в третью урну. Найти вероятность того, что шар, вынутый наудачу из третьей урны, окажется белым.
2. В коробке 5 новых и 3 уже использованных теннисных мяча. Для игры из коробки наудачу берут два мяча и затем возвращают их в коробку. Найти вероятность того, что при случайном выборе двух мячей из этой коробки для следующей игры оба мяча окажутся новыми.
3. Есть пять кубиков с цифрами на гранях $1, 2, \ldots, 6$ и два кубика, у которого на двух гранях цифра 1, на двух других — цифра 2, а на оставшихся — цифра 3. Наудачу выбирают один из кубиков и бросают. Какова вероятность того, что выпадет четное число?
4. Три стрелка одновременно и независимо стреляют по кабану. Известно, что вероятность попадания 1-го равна $0,8$, 2-го — $0,6$, а 3-го — $0,2$. В результате кабан убит и в нем обнаружено две пули. Для каждого из стрелков найти вероятность того, что он не попал.
5. По каналу связи передается одна из последовательностей букв AAAA, BBBB, CCCC с вероятностями $1/2$, $1/3$ и $1/6$ соответственно. Каждая передаваемая буква принимается правильно с вероятностью $0,8$ и с вероятностями $0,1$ и $0,1$ принимается за каждую из двух других букв. Предполагается, что буквы искажаются независимо друг от друга. Какова вероятность того, что было передано AAAA, если принято ABCA?
6. Некоторым заболеванием болеет 1 процент населения. Для диагностирования используется тест, надежность которого 90 процентов (если человек болен, то вероятность определения этого тестом равна 0,9; если он здоров, вероятность того, что тест это подтвердит, равна 0,9). Какова вероятность того, что человек действительно болен, если тест показывает, что он болен?

Глава 4. Схема Бернулли и предельные теоремы для нее

Схемой Бернулли называют математическую модель серии из n одинаковых независимых экспериментов (их принято называть ***испытаниями***), в каждом из которых может произойти (или не произойти) одно и то же событие, условно называемое ***успехом***, вероятность которого постоянна и равна p (***вероятность успеха***). Противоположное событие называют ***неудачей***. Ее вероятность обозначают q. Очевидно, $q = 1 - p$. Примерами случайных экспериментов, описываемых схемой Бернулли, являются:

- n-кратное бросание монеты (испытание — бросок, успех — выпадение герба, вероятность успеха — $p = \dfrac{1}{2}$, если монета правильная (симметричная и однородная), или $p \neq \dfrac{1}{2}$, если монета неправильная);
- n-кратное бросание игральной кости (испытание — бросок, если успехом считать, например, выпадение шестерки, вероятность успеха — $p = \dfrac{1}{6}$);
- n-кратная стрельба по мишени, при условии, что выстрелы независимые (испытание — выстрел, успех — попадание, p — вероятность попадания).

4.1. Основные формулы схемы Бернулли

Будем обозначать через $P_n(m)$ вероятность того, что в серии из n испытаний Бернулли число успехов равно m $(0 \le m \le n)$.

Формула Бернулли:

$$P_n(m) = C_n^m p^m q^{n-m}\,. \tag{4.1}$$

Доказательство. Рассмотрим серию из n испытаний Бернулли с вероятностью успеха p и следующие события:

$$A = \{\text{произошло } m \text{ успехов}\}\,,$$

$$A_i = \{i\text{-е испытание закончилось успехом}\}\,.$$

Событие A представляет собой сумму произведений событий следующего вида:

$$\begin{aligned}A = {} & A_1A_2\ldots A_m\overline{A}_{m+1}\overline{A}_{m+2}\ldots\overline{A}_n + \\ & + A_1A_2\ldots A_{m-1}\overline{A}_mA_{m+1}\overline{A}_{m+2}\ldots\overline{A}_n + \ldots \\ & \cdots + \overline{A}_1\overline{A}_2\ldots\overline{A}_{n-m}A_{n-m+1}\ldots A_n\,.\end{aligned}$$

Каждое слагаемое в правой части представляет собой один из возможных способов протекания серии из n испытаний Бернулли, при котором происходит ровно m успехов. Оно содержит n множителей, где i-й множитель — результат i-го испытания (A_i — успех, $\overline{A}_i$ — неудача). В каждом произведении m успехов и $n-m$ неудач. Очевидно, слагаемые представляют собой попарно несовместные события, поэтому вероятность их суммы равна сумме их вероятностей. При этом каждое из них представляет собой произведение независимых событий, а значит вероятность каждого из произведений равна произведению вероятностей входящих в него событий, т. е. p^mq^{n-m}. Для каждого слагаемого номера испытаний, закончившихся успехами, образуют сочетание из n номеров по m, поэтому количество слагаемых равно C_n^m. В результате $P(A) = P_n(m) = C_n^m p^m q^{n-m}$. ■

Обозначим через $P_n(m_1, m_2)$ вероятность того, что в серии из n испытаний Бернулли число успехов не меньше m_1 и не больше m_2 $(0 \le m_1 < m_2 \le n)$. Для этой вероятности справедливы равенства

$$
\boxed{\begin{aligned}
P_n(m_1, m_2) &= \sum_{m=m_1}^{m_2} C_n^m p^m q^{n-m} = \\
&= 1 - \sum_{m=0}^{m_1-1} C_n^m p^m q^{n-m} - \sum_{m=m_2+1}^{n} C_n^m p^m q^{n-m} .
\end{aligned} \qquad (4.2)}
$$

Доказательство. Рассмотрим серию из n испытаний Бернулли и следующие события:

$$
A = \{\text{число успехов не меньше } m_1 \text{ и не больше } m_2\} ,
$$
$$
A_m = \{\text{число успехов равно } m\} .
$$

Очевидно, $A = \sum\limits_{m=m_1}^{m_2} A_m$, причем слагаемые представляют собой попарно несовместные события, так что вероятность суммы этих событий равна сумме их вероятностей. Отсюда и из формулы Бернулли (4.1) следует первое равенство формулы (4.2). С другой стороны,

$$
A = \overline{\sum_{m=0}^{m_1-1} A_m + \sum_{m=m_2+1}^{n} A_m} ,
$$

откуда по свойству (Р-6), снова используя попарную несовместность событий A_m и формулу Бернулли (4.1), получаем второе равенство формулы (4.2). ■

Ⓡ Второе равенство формулы (4.2) можно вывести и с помощью известного разложения ***бинома Ньютона***:

$$
\sum_{m=0}^{n} C_n^m p^m q^{n-m} = (p+q)^n = 1 .
$$

Задача 4.1.1. Стрелок делает семь выстрелов по мишени. Найти вероятность того, что число попаданий будет от четырех до шести, если выстрелы независимы и вероятность попадания при каждом равна $\dfrac{2}{3}$.

Решение. Перед нами схема Бернулли (выстрел — испытание, попадание — успех, $= \frac{2}{3}$, $n = 7$). Искомая вероятность равна

$$P_7(4,6) = C_7^4\left(\frac{2}{3}\right)^4\left(\frac{1}{3}\right)^3 + C_7^5\left(\frac{2}{3}\right)^5\left(\frac{1}{3}\right)^2 + C_7^6\left(\frac{2}{3}\right)^6\left(\frac{1}{3}\right) =$$

$$= \frac{35\cdot 16 + 21\cdot 32 + 7\cdot 64}{3^7} = \frac{1680}{2187} \approx 0,768\,. \qquad ■$$

Задача 4.1.2. В кармане курящего человека лежат две коробки спичек по 100 штук в каждой. Каждый раз, закуривая, он достает одну из коробок, извлекая их наудачу. Найти вероятность того, что когда одна из коробок опустеет, в другой будет ровно m спичек.

Решение. Случайный эксперимент, о котором идет речь в задаче, представляет собой серию испытаний Бернулли, где испытанием является выбор коробки для извлечения очередной спички. Успехом при этом будем считать выбор той коробки, которая в конце эксперимента оказывается пустой. Тогда вероятность успеха $p = \dfrac{1}{2}$, а число испытаний $n = 200-m$. В результате искомая вероятность равна

$$P_{200-m}(100) = C_{200-m}^{100}\left(\frac{1}{2}\right)^{200-m}. \qquad ■$$

4.2. Предельные теоремы для схемы Бернулли

Вычисления вероятностей с помощью формулы (4.1), а особенно с помощью формулы (4.2), становятся очень трудоемкими при больших значениях n из-за необходимости вычислять факториалы больших чисел. В этом разделе мы рассмотрим теоремы, дающие возможность использовать в такой ситуации более простые приближенные формулы.

Теорема 4.2.1. — Пуассон.

$$P_n(m) \xrightarrow[\substack{n\to\infty \\ np=\lambda}]{} \frac{\lambda^m}{m!}e^{-\lambda}, \quad \lambda \in \mathbb{R}\,. \tag{4.3}$$

Доказательство. Преобразуем выражение для $P_n(m)$ по формуле Бернулли, полагая $p = \dfrac{\lambda}{n}$:

$$P_n(m) = C_n^m p^m q^{n-m} = \frac{n!}{m!(n-m)!}\left(\frac{\lambda}{n}\right)^m \left(1-\frac{\lambda}{n}\right)^{(n-m)} =$$
$$= \frac{\lambda^m}{m!}\cdot\frac{n(n-1)\dots(n-m+1)}{n^m}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-m} =$$
$$= \frac{\lambda^m}{m!}\underbrace{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{m-1}{n}\right)}_{\downarrow\atop 1}\left(1-\frac{\lambda}{n}\right)^n\underbrace{\left(1-\frac{\lambda}{n}\right)^{-m}}_{\downarrow\atop 1}.$$

Множители, отмеченные фигурными скобками, при $n \to \infty$ стремятся к 1, поскольку $m = \mathrm{Const}$, а из второго замечательного предела следует $\lim\limits_{n\to\infty}\left(1-\dfrac{\lambda}{n}\right)^n = e^{-\lambda}$. ■

Теорема Пуассона дает нам следующую приближенную формулу для вычисления вероятностей в схеме Бернулли:

$$\boxed{P_n(m) \approx \frac{\lambda^m}{m!}e^{-\lambda}, \quad \lambda = np. \qquad (4.4)}$$

(R) Из утверждения теоремы Пуассона следует, что формула (4.4) дает хорошие приближения при больших значениях n, если при этом вероятность успеха p очень мала, так, что величина $\lambda = np$ принимает «разумное» значение (на практике ее применяют при $npq < 9$).

Задача 4.2.1. Пекарня выпекает 1000 булочек с изюмом. Сколько надо положить в тесто изюма, чтобы вероятность найти в произвольно взятой булочке хотя бы одну изюминку была не менее $0,95$?

Решение. Случайный эксперимент, о котором идет речь в задаче, состоит в распределении изюминок по булочкам. Будем смотреть на него как на схему Бернулли, где испытание — выбор булочки для очередной изюминки. Обозначим через n неизвестное число испытаний, т. е. изюминок. Пусть для проверки выбрана некоторая булочка. Успехом в испытании будем считать попадание изюминки в эту булочку. Таким образом, вероятность успеха $p = 0,001$. Вероятность того, что в булочке, выбранной для проверки, будет хотя бы одна изюминка, равна вероятности того, что в описанной схеме Бернулли число успехов больше либо равно 1. Находим неизвестное число n из условия

$$P_n(1, n) = 1 - P_n(0) \geq 0,95 \;\Leftrightarrow\; P_n(0) \leq 0,05\,.$$

По формуле Бернулли $P_n(0) = q^n = 0,999^n$, так что для n получаем требование

$$0,999^n \leq 0,05 \;\Leftrightarrow\; n \geq \log_{0,999} 0,05 = \frac{\ln 0,05}{\ln 0,999} \approx 2994,23\,.$$

Таким образом, для выполнения условия требуется 2995 изюминок.

Рассмотрим результат, получающийся с помощью приближенной формулы (4.4). Потребуем

$$P_n(0) \approx \frac{\lambda^0}{0!} e^{-\lambda} = e^{-\lambda} \leq 0,05 \;\Leftrightarrow\; \lambda \geq -\ln 0,05 = \ln 20 \approx 2,9957\,.$$

В нашем случае $\lambda = n \cdot 0,001$, поэтому приходим к требованию $n \geq 2955,7$. Таким образом, относительная погрешность приближения для результата, полученного с помощью приближенной формулы, меньше 10^{-3}. ■

Следующая теорема описывает асимптотику вероятности $P_n(m)$ при $n \to \infty$ и фиксированной вероятности успеха p. Ее доказательство опирается на формулу Стирлинга, которая описывает асимптотику факториала при $n \to \infty$.

Формула Стирлинга:

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n, \quad n \to \infty\,. \tag{4.5}$$

Рассмотрим достаточно грубую оценку, проясняющую устройство правой части формулы Стирлинга. Для этого рассмотрим логарифм факториала:

$$\ln n! = \sum_{k=1}^{n} \ln k = \frac{\ln 1 + \ln 2}{2} + \frac{\ln 2 + \ln 3}{2} + \ldots$$
$$\ldots + \frac{\ln(n-1) + \ln n}{2} + \frac{\ln n}{2} = \sum_{k=2}^{n} \frac{\ln(k-1) + \ln k}{2} + \frac{\ln n}{2}.$$

Величина $\dfrac{\ln(k-1)+\ln k}{2}$ равна площади трапеции, вписанной в область, ограниченную графиком функции $\ln x$ на отрезке $[k-1;k]$. Ее основания параллельны оси OY, их длины равны $\ln(k-1)$ и $\ln k$, а длина высоты равна 1. В результате сумма этих величин равна площади многоугольника, представляющего собой объединение таких трапеций по k от 2 до n (см. рис. 4.1). Площадь этого многоугольника не превосходит площади, заключенной под графиком логарифма на отрезке $[1;n]$, поэтому справедлива оценка

$$\ln n! \leq \int_1^n \ln x\,dx + \frac{\ln n}{2} = x\ln x\Big|_1^n - \int_1^n x\,d\ln x + \frac{\ln n}{2} =$$
$$= n\ln n - n + 1 + \ln\sqrt{n} = \ln\left[e\sqrt{n}\left(\frac{n}{e}\right)^n\right].$$

Таким образом, получаем неравенство

$$n! \leq \sqrt{e^2 n}\left(\frac{n}{e}\right)^n. \tag{4.6}$$

Аналогичная по порядку оценка для факториала может быть получена и снизу, например,

$$n! \geq \sqrt{\frac{8e^2 n}{27}}\left(\frac{n}{e}\right)^n. \tag{4.7}$$

Из оценок (4.6) и (4.7) следует, что для некоторой константы C справедливо соотношение

$$n! \sim C\sqrt{n}\left(\frac{n}{e}\right)^n.$$

Формула Стирлинга утверждает, что $C = \sqrt{2\pi}$.

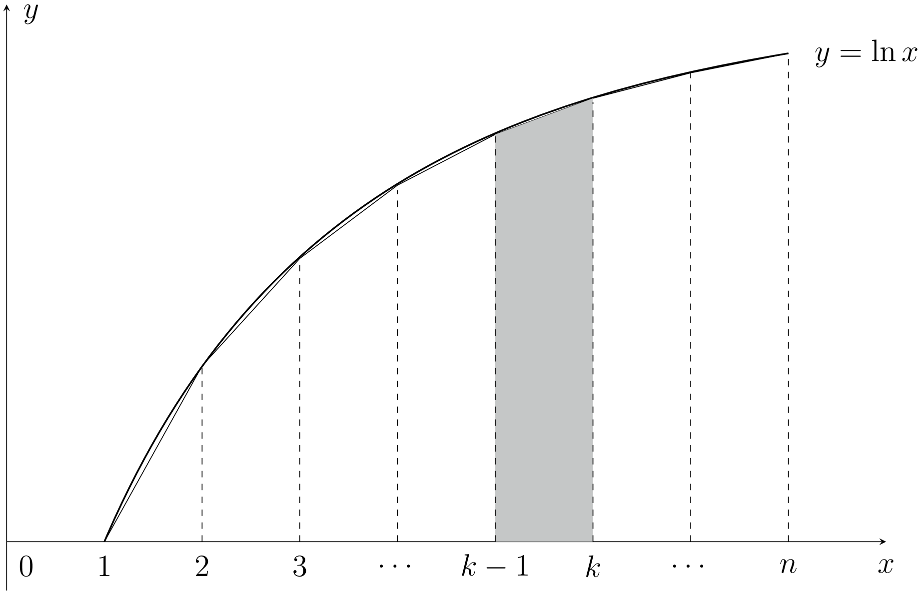


Рис. 4.1. Оценка $\ln n!$

Теорема 4.2.2. — Муавр–Лаплас (локальная).

$$\sqrt{npq}\,P_n(m) \sim \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\,, \text{ при } n, m \to \infty\,, \tag{4.8}$$

если $x = x_n(m) = \dfrac{m - np}{\sqrt{npq}} \in [a, b]$, где $a, b \in \mathbb{R}$ — некоторые константы.

Доказательство. Пользуясь формулой Стирлинга, заменим факториалы в выражении $P_n(m)$ по формуле Бернулли эквивалентными бесконечно большими при $n, m \to \infty$:

$$\sqrt{npq}\,P_n(m) = \sqrt{npq}\frac{n!}{m!(n-m)!}p^m q^{n-m} \sim$$
$$\sim \sqrt{npq}\frac{\sqrt{2\pi n}\,n^n e^m e^{n-m} p^m q^{n-m}}{e^n\sqrt{2\pi m}\,m^m\sqrt{2\pi(n-m)}\,(n-m)^{n-m}}\,.$$

Сокращая общие множители и группируя оставшиеся, получим

$$\sqrt{npq}\,P_n(m) \sim$$
$$\sim \frac{1}{\sqrt{2\pi}}\left(\frac{m}{np}\right)^{-m}\left(\frac{n-m}{nq}\right)^{-(n-m)}\left(\frac{m}{np}\right)^{-\frac{1}{2}}\left(\frac{n-m}{nq}\right)^{-\frac{1}{2}}\,. \tag{4.9}$$

По условию теоремы при $n, m \to \infty$ величина $x = \dfrac{m - np}{\sqrt{npq}}$ остается ограниченной. Выразим через нее m и $n - m$:

$$m = np + x\sqrt{npq}\,, \quad n - m = nq - x\sqrt{npq}\,.$$

Подставив эти выражения в правую часть асимптотического соотношения (4.9), получим:

$$\sqrt{npq}\,P_n(m) \sim$$
$$\sim \frac{1}{\sqrt{2\pi}}\left(1 + x\sqrt{\frac{q}{np}}\right)^{-(np+x\sqrt{npq})}\left(1 - x\sqrt{\frac{p}{nq}}\right)^{-(nq-x\sqrt{npq})}\cdot$$
$$\cdot\left(1 + x\sqrt{\frac{q}{np}}\right)^{-\frac{1}{2}}\left(1 - x\sqrt{\frac{p}{nq}}\right)^{-\frac{1}{2}} \sim$$
$$\sim \frac{1}{\sqrt{2\pi}}\exp\left[-(np + x\sqrt{npq})\ln\left(1 + x\sqrt{\frac{q}{np}}\right) -\right.$$
$$\left. - (nq - x\sqrt{npq})\ln\left(1 - x\sqrt{\frac{p}{nq}}\right)\right],$$

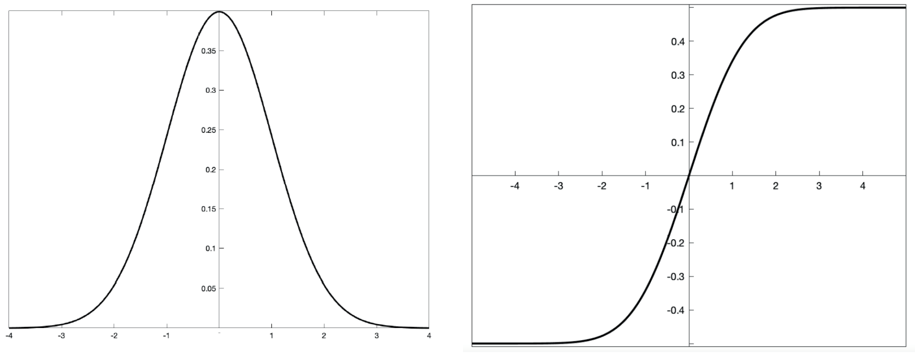
поскольку $\left(1 + x\sqrt{\dfrac{q}{np}}\right)^{-\frac{1}{2}}\left(1 - x\sqrt{\dfrac{p}{nq}}\right)^{-\frac{1}{2}} \to 1$ при $n \to \infty$.

Используя разложение логарифма по формуле Тейлора–Маклорена,

$$\ln(1 + x) = x - x^2/2 + o(x^2)\,, \text{ при } x \to 0\,,$$

получим:

$$\sqrt{npq}\,P_n(m) \sim \frac{1}{\sqrt{2\pi}}\exp\left[-(np + x\sqrt{npq})\left(x\sqrt{\frac{q}{np}} - \frac{x^2 q}{2np} + o\left(\frac{1}{n}\right)\right) -\right.$$
$$\left. - (nq - x\sqrt{npq})\left(-x\sqrt{\frac{p}{nq}} - \frac{x^2 p}{2nq} + o\left(\frac{1}{n}\right)\right)\right] =$$
$$= \frac{1}{\sqrt{2\pi}}\exp\left[-x\sqrt{npq} + \frac{x^2 q}{2} - x^2 q + x\sqrt{npq} + \frac{x^2 p}{2} - x^2 p + o(1)\right] =$$
$$= \frac{1}{\sqrt{2\pi}}\exp\left[-\frac{x^2(p + q)}{2} + o(1)\right] \sim \frac{1}{\sqrt{2\pi}}\exp\left[-\frac{x^2}{2}\right]. \qquad \blacksquare$$



а) б)

Рис. 4.2. Графики а) $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, б) $\Phi_0(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int\limits_0^x e^{-\frac{t^2}{2}} dt$

Функция в правой части асимптотической формулы (4.8), определенная равенством

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}, \tag{4.10}$$

называется ***функцией Гаусса***. Это четная строго положительная функция, стремящаяся к нулю на бесконечности. Ее график имеет характерную форму колокола (см. рис. 4.2., а). При этом площадь под графиком равна 1. Это с помощью замены переменной можно получить, используя известный ***интеграл Пуассона***:

$$\int\limits_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}\,. \tag{4.11}$$

Из локальной теоремы Муавра–Лапласа следует приближенная формула, которую можно использовать для вычисления вероятностей в схеме Бернулли при больших значениях n и «не слишком маленьких» p и q.

Локальная формула Муавра–Лапласа:

$$P_n(m) \approx \varphi\big(x_n(m)\big) \frac{1}{\sqrt{npq}}\,, \text{ где } \; x_n(m) = \frac{m - np}{\sqrt{npq}}\,. \tag{4.12}$$

Однако при больших значениях n вероятности $P_n(m)$ очень малы, а практический интерес представляют их суммы, через которые выражаются вероятности $P_n(m_1, m_2)$ (см. формулу (4.2)). Следующая теорема описывает асимптотику этих вероятностей.

Теорема 4.2.3. (Муавр–Лаплас, интегральная)

$$P_n(m_1, m_2) \sim \frac{1}{\sqrt{2\pi}} \int\limits_a^b e^{-\frac{x^2}{2}} dx\,, \quad \text{при } n \to \infty\,, \tag{4.13}$$

если $\dfrac{m_1 - np}{\sqrt{npq}} \to a\,,\ \dfrac{m_2 - np}{\sqrt{npq}} \to b\,,\ a, b \in \mathbb{R}$.

Доказательство. Ограничимся нестрогим обоснованием. Используя формулу (4.2), а затем асимптотическую формулу (4.8), получаем:

$$\begin{aligned} P_n(m_1, m_2) &= \sum_{m=m_1}^{m_2} P_n(m) \sim \sum_{m=m_1}^{m_2} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x_n^2(m)}{2}} \frac{1}{\sqrt{npq}} = \\ &= \sum_{m=m_1}^{m_2} \varphi\big(x_n(m)\big)\Delta x_m\,. \end{aligned} \tag{4.14}$$

Точки $x_n(m) = \dfrac{m - np}{\sqrt{npq}}$, где $m = m_1, m_1 + 1, \ldots, m_2$, образуют равномерное разбиение отрезка $\left[\dfrac{m_1 - np}{\sqrt{npq}}; \dfrac{m_2 - np}{\sqrt{npq}}\right]$ на отрезки длины $\Delta x_m = x_n(m+1) - x_n(m) = \dfrac{1}{\sqrt{npq}}$, так что сумма в правой части равенства (4.14) представляет собой интегральную сумму для интеграла от функции Гаусса, построенную по этому разбиению. При $n \to \infty$ диаметр разбиения, равный $\dfrac{1}{\sqrt{npq}}$, стремится к нулю, при этом, по условию теоремы, отрезок $\left[\dfrac{m_1 - np}{\sqrt{npq}}; \dfrac{m_2 - np}{\sqrt{npq}}\right]$ стремится к $[a, b]$, так что предел интегральных сумм равен интегралу из правой части формулы (4.13). ■

R Первообразная функции Гаусса не выражается через элементарные функции, так что интеграл в правой части асимптотической формулы (4.13) — «неберущийся». При практической работе с ним используют следующие специальные функции:

$$\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-\frac{t^2}{2}} dt\,, \tag{4.15}$$

$$\Phi(x) = \sqrt{\frac{2}{\pi}} \int\limits_0^x e^{-\frac{t^2}{2}} dt\,, \tag{4.16}$$

$$\mathrm{Erf}\,(x) = \frac{2}{\sqrt{\pi}} \int\limits_0^x e^{-t^2} dt\,. \tag{4.17}$$

Первая из них — функция $\Phi_0(x)$ — одна из первообразных функции Гаусса. Это нечетная строго возрастающая на всей оси функция, стремящаяся к $\frac{1}{2}$ при $x \to +\infty$ и к $-\frac{1}{2}$ при $x \to -\infty$ (см. рис. 4.2., б). Функция $\Phi(x)$, определенная формулой (4.16), — ***функция Лапласа***. Функцию $\mathrm{Erf}\,(x)$ называют ***функцией ошибок***. Эти три функции связаны друг с другом соотношениями

$$\Phi_0(x) = \frac{1}{2}\Phi(x) = \frac{1}{2}\,\mathrm{Erf}\,\left(\frac{x}{\sqrt{2}}\right)\,.$$

Значения этих и обратных к ним функций можно вычислять в пакетах прикладных программ, таких как *MatLab*, *MatCad*, *Maple* и др. Кроме того, во многих учебниках по теории вероятностей можно найти таблицы значений этих функций.

Из утверждения интегральной теоремы Муавра–Лапласа, с помощью формулы Ньютона–Лейбница, получается следующая приближенная формула:

Интегральная формула Муавра–Лапласа:

$$P_n(m_1, m_2) \approx \frac{1}{\sqrt{2\pi}} \int\limits_a^b e^{-\frac{x^2}{2}} dx = \Phi_0(x)\Big|_a^b = \\ = \frac{1}{2}\Phi(x)\Big|_a^b = \frac{1}{2}\,\mathrm{Erf}\,\left(\frac{x}{\sqrt{2}}\right)\Big|_a^b\,, \tag{4.18}$$

где $a = \dfrac{m_1 - np}{\sqrt{npq}}\,,\ b = \dfrac{m_2 - np}{\sqrt{npq}}\,.$

Задача 4.2.2. Игральную кость бросают 700 раз. Найти вероятность того, что шестерка выпадет не менее 100 и не более 120 раз.

Решение. В задаче рассматривается серия испытаний Бернулли, где число испытаний (бросков) $n = 700$, успехом в испытании считается выпадение шестерки, $p = \dfrac{1}{6}$, $q = \dfrac{5}{6}$. По формуле (4.18) найдем искомую вероятность:

$$P_{700}(100, 120) \approx \frac{1}{2}\big(\Phi(b) - \Phi(a)\big)\,,$$

где $\Phi(x)$ — функция Лапласа,

$$a = \frac{100 - 700 \cdot \frac{1}{6}}{\sqrt{700 \cdot \frac{1}{6} \cdot \frac{5}{6}}} \approx -1,69\,, \quad b = \frac{120 - 700 \cdot \frac{1}{6}}{\sqrt{700 \cdot \frac{1}{6} \cdot \frac{5}{6}}} \approx 0,34\,.$$

Используя приближенные значения

$$\Phi(1,69) \ \approx \ 0,90897\,, \quad \Phi(0,34) \ \approx \ 0,26614$$

и пользуясь нечетностью функции Лапласа, получаем:

$$P_{700}(100, 120) \approx \frac{1}{2}(0,26614 + 0,90897) = 0,587556\,.$$

■

Задача 4.2.3. Город ежедневно посещают 1500 туристов. Каждый из них выбирает для обеда один из трех городских ресторанов с равными вероятностями и независимо друг от друга. Владелец одного из ресторанов хочет, чтобы с вероятностью 0,99 все туристы, пришедшие в его ресторан, смогли там пообедать одновременно. Сколько мест должно быть для этого в ресторане?

Решение. Здесь мы имеем дело со схемой Бернулли, где испытанием является выбор туристом ресторана для обеда. Успехом в таком испытании будем считать выбор ресторана, принадлежащего данному владельцу. По условию испытания являются независимыми, их число $n = 1500$, вероятность успеха $p = \dfrac{1}{3}$. Пусть x — искомое число мест в ресторане. Рассмотрим событие

$$A = \{\text{в ресторан пришло не более } x \text{ туристов}\}\,.$$

Наступление события A означает, что всем пришедшим в ресторан туристам хватило мест. Потребуем выполнения условия

$$P(A) = P_{1500}(0, x) \geq 0,99\,.$$

По формуле (4.18) это эквивалентно требованию

$$\frac{1}{2}\left(\Phi(b) - \Phi(a)\right) \geq 0,99\,, \tag{4.19}$$

где $a = \dfrac{0 - 1500 \cdot \dfrac{1}{3}}{\sqrt{1500 \cdot \dfrac{1}{3} \cdot \dfrac{2}{3}}} \approx -27,38$, $b = \dfrac{x - 1500 \cdot \dfrac{1}{3}}{\sqrt{1500 \cdot \dfrac{1}{3} \cdot \dfrac{2}{3}}}$. При $x > 5$ можно считать $\Phi(x) \approx 1$, поэтому, в силу нечетности функции Лапласа, $\Phi(a) \approx -1$. В результате требование (4.19) принимает вид

$$\Phi(b) \geq 0,98\,.$$

С помощью таблиц значений функции Лапласа, или пользуясь пакетом прикладных программ, находим минимальное b, удовлетворяющее этому условию:

$$b = \Phi^{-1}(0,98) \approx 2,33\,.$$

Решая уравнение

$$\frac{x - 1500 \cdot \dfrac{1}{3}}{\sqrt{1500 \cdot \dfrac{1}{3} \cdot \dfrac{2}{3}}} = 2,33\,,$$

находим $x = 1500 \cdot \dfrac{1}{3} + 2,33 \cdot \sqrt{1500 \cdot \dfrac{1}{3} \cdot \dfrac{2}{3}} \approx 542,33$. Следовательно, для того чтобы с вероятностью не меньшей $0,99$ всем, пришедшим в ресторан, хватило там мест, в нем должно быть, как минимум, 543 места. ■

4.3. Задачи для самостоятельного решения

1. Чему равно наиболее вероятное число успехов при n испытаниях Бернулли с вероятностью успеха p? Найти m_0 — наиболее вероятное число попаданий при шести выстрелах, если вероятность попадания при одном выстреле равна $0,4$. Какова вероятность того, что число попаданий будет не меньше $m_0 - 1$ и не больше, чем $m_0 + 1$?
2. Книга, состоящая из 500 страниц, содержит 800 опечаток. Найти вероятность того, что на произвольно взятой странице не менее двух опечаток.
3. Симметричная монета брошена 500 раз. Найти границы, в которых число выпаданий герба будет заключено с вероятностью 0,97.
4. В сборочном цехе 250 рабочих. Каждый из них потребляет электроэнергию в среднем 12 минут в течение часа. Какой должна быть минимальная мощность силовой установки цеха, чтобы вероятность перегрузки в произвольный момент времени была не больше 0,0001? Под мощностью силовой установки понимается максимальное число рабочих, которое она может снабжать энергией одновременно.
5. Вероятность того, что при сдаче карт для игры в покер игрок получит десятку, валета, даму, короля и туза одной масти, равна $1/649740$. Сколько надо сыграть игр, чтобы вероятность не получить ни одного такого набора карт была не больше $1/e \approx 1/3$?
6. К станции подходит электропоезд из восьми вагонов. В каждом вагоне 72 сидячих места. На станции поезд ожидают 600 пассажиров, из которых каждый с равной вероятностью, независимо от других может сесть в любой вагон. Найти вероятность того, что в произвольно взятом вагоне после посадки останется не менее двух свободных мест.

7. Для определения доли женщин в некотором обществе производится случайная выборка. Каким должен быть объем выборки, чтобы с вероятностью 0,95 доля женщин в ней отличалась от истинного значения менее, чем на 0,01?

Глава 5. Случайные величины

При моделировании случайных явлений интерес, как правило, представляют их всевозможные числовые характеристики. В теории вероятностей это приводит к такому понятию, как ***случайная величина***.

Пусть вероятностное пространство $(\Omega, \mathcal{F}, P)$ соответствует некоторому случайному эксперименту. В качестве примера можно представить себе приземление самолета на аэродроме. Каждый новый заходящий на посадку самолет — новый случайный эксперимент. В такой ситуации можно считать, что каждый исход $\omega \in \Omega$ представляет собой один из множества всевозможных сценариев посадки, а сигма-алгебра событий $\mathcal{F}$ содержит все события, которые при этом могут произойти (или не произойти). Можно представить себе огромное множество числовых характеристик рассматриваемого случайного явления — от веса самолета, вертикальной и горизонтальной составляющих его скорости, температуры масла в двигателях и температуры воздуха на борту до количества пассажиров на борту, их общего веса и пульса пилота в момент соприкосновения самолета с землей. Каждый из этих показателей является случайной величиной, в том смысле, что для каждого $\omega \in \Omega$ он принимает, вообще говоря, свое значение. Таким образом, мы приходим к понятию случайной величины как функции от ω — исхода случайного эксперимента.

При этом оказывается, что для построения содержательной теории, случайной величиной следует считать не всякую функцию от $\omega \in \Omega$. Рассмотрим понятие измеримости, которое является ключевым в строгом определении понятия «случайная величина».

5.1. Измеримые функции

Определение 5.1.1. Пусть $(X, \mathcal{X})$ и $(Y, \mathcal{Y})$ — ***измеримые пространства***[a], функция $f : X \to Y$ называется $\mathcal{X}|\mathcal{Y}$-***измеримой*** (обозначение: $f \in \mathcal{X}|\mathcal{Y}$), если

$$\forall\ B \in \mathcal{Y} \quad f^{-1}(B) = \{x \in X \mid f(x) \in B\} \in \mathcal{X}\,,$$

т. е. прообраз любого множества B из сигма-алгебры $\mathcal{Y}$ является множеством, принадлежащим сигма-алгебре $\mathcal{X}$.

[a] ***Измеримым пространством*** называется пара $(X\mathcal{X})$, где X — произвольное непустое множество, а $\mathcal{X}$ — некоторая сигма-алгебра его подмножеств. Если на $\mathcal{X}$ при этом задана сигма-аддитивная мера μ, то тройка $(X, \mathcal{X}, \mu)$ называется ***пространством с мерой***. Таким образом, вероятностное пространство $(\Omega, \mathcal{F}, P)$— это частный случай пространства с мерой, а пара $(\Omega, \mathcal{F})$ — измеримое пространство.

Если в роли множества Y выступает $\mathbb{R}$, то в роли сигма-алгебры $\mathcal{Y}$ всегда используют $\mathcal{B}(\mathbb{R})$ — ***борелевскую*** сигма-алгебру подмножеств $\mathbb{R}$, а $\mathcal{X}|\mathcal{B}(\mathbb{R})$-измеримую функцию называют просто ***измеримой***.

Определение 5.1.2. Пусть X — непустое множество, а $\mathcal{K}$ — некоторая совокупность его подмножеств. Минимальная сигма-алгебра подмножеств X, содержащая $\mathcal{K}$, называется сигма-алгеброй, ***порожденной*** системой $\mathcal{K}$. Ее обозначают символом $\sigma(\mathcal{K})$.

▪ Пример 5.1.1. Пусть $\mathcal{K} = \big\{(-\infty; 0)\big\}$. Тогда

$$\sigma(\mathcal{K}) = \big\{(-\infty; 0), [0; +\infty), \varnothing, \mathbb{R}\big\}\,.$$

Действительно, вместе с промежутком $(-\infty; 0)$ содержащая его сигма-алгебра должна содержать дополнение до него, т. е. промежуток $[0; +\infty)$, объединение этих промежутков — $\mathbb{R}$, и пустое множество. Совокупность этих четырех множеств, очевидно, является минимальной сигма-алгеброй, содержащей $\mathcal{K}$. ▪

Определение 5.1.3. $\mathcal{B}(\mathbb{R})$ — ***борелевская*** сигма-алгебра подмножеств $\mathbb{R}$ (или сигма-алгебра борелевских подмножеств $\mathbb{R}$) — сигма-алгебра, порожденная совокупностью всех промежутков в $\mathbb{R}$, т. е.

$$\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{P}),$$

где $\mathcal{P}$ — совокупность всех промежутков в $\mathbb{R}$.

Борелевская сигма-алгебра подмножеств множества $\mathbb{R}$ не совпадает с совокупностью всех подмножеств $\mathbb{R}$, тем не менее, это очень богатая система. Она по определению содержит все промежутки, а так как представляет собой сигма-алгебру, в нее входят и всевозможные конечные и счетные объединения промежутков. В частности борелевскими являются одноэлементные множества, так как они являются разностями промежутков: $\{a\} = [a; b) \setminus (a; b)$. Как следствие, любые конечные и счетные множества являются борелевскими, так как представляют собой конечные или счетные объединения борелевских множеств. Таким образом, борелевскими являются те подмножества $\mathbb{R}$, которыми мы обычно пользуемся. Конечно, ими $\mathcal{B}(\mathbb{R})$ не исчерпывается и среди борелевских есть множества очень сложной структуры, так что не существует конструктивного описания всех борелевских подмножеств $\mathbb{R}$. Еще одно описание сигма-алгебры борелевских подмножеств $\mathbb{R}$ дает следующее утверждение.

Предложение 5.1.1. $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{P}_0)$, где $\mathcal{P}_0 = \{(-\infty; c) \,|\, c \in \mathbb{R}\}$.

Доказательство. Требуется доказать, что $\sigma(\mathcal{P}_0) = \sigma(\mathcal{P})$. Действительно, $\mathcal{P}_0 \subset \mathcal{P}$. Из этого следует, что $\mathcal{P}_0 \subset \sigma(\mathcal{P})$. Поскольку $\sigma(\mathcal{P}_0)$ — минимальная сигма-алгебра, содержащая $\mathcal{P}_0$, из этого вложения следует, что она содержится в $\sigma(\mathcal{P})$:

$$\sigma(\mathcal{P}_0) \subseteq \sigma(\mathcal{P}). \tag{5.1}$$

Для доказательства обратного вложения заметим, что $\mathcal{P} \subseteq \sigma(\mathcal{P}_0)$. Действительно, любой промежуток в $\mathbb{R}$ можно построить из промежутков вида $(-\infty; c)$ с помощью не более чем счетного количества теоретико-множественных операций:

$$[c; +\infty) = \mathbb{R} \setminus (-\infty; c),$$

$$
\begin{aligned}
&(-\infty; c] = \bigcap_{k=1}^{\infty} (-\infty; c + \frac{1}{k})\,, \quad (c; +\infty) = \mathbb{R} \setminus (-\infty; c]\,,\\
&(a; b) = (-\infty; b) \setminus (-\infty; a]\,,\\
&(a; b] = (-\infty; b] \setminus (-\infty; a]\,,\\
&[a; b) = (-\infty; b) \setminus (-\infty; a)\,,\\
&[a; b] = (-\infty; b] \setminus (-\infty; a)\,,\\
&\mathbb{R} = \bigcup_{k=1}^{\infty} (-\infty; k)\,.
\end{aligned}
$$

Значит любой промежуток в $\mathbb{R}$ принадлежит сигма-алгебре, порожденной системой $\mathcal{P}_0$. В результате получаем

$$
\sigma(\mathcal{P}) \subseteq \sigma(\mathcal{P}_0) \tag{5.2}
$$

как минимальная σ-алгебра, содержащая систему $\mathcal{P}$. Из вложений (5.1) и (5.2) следует доказываемое равенство. ■

Оказывается, для проверки измеримости функции $f : X \to Y$ относительно сигма-алгебр $\mathcal{X}$ и $\mathcal{Y}$ не обязательно проверять условие $f^{-1}(B) \in \mathcal{X}$ для всех $B \in \mathcal{Y}$. Об этом говорит следующее предложение.

Предложение 5.1.2. Пусть $(X, \mathcal{X})$ — измеримое пространство, $f : X \to Y$, $\mathcal{K}$ — некоторая совокупность подмножеств множества Y. Если

$$
\forall\ B \in \mathcal{K} \quad f^{-1}(B) \in \mathcal{X}\,,
$$

то $f \in \mathcal{X}|\sigma(\mathcal{K})$.

Доказательство. Докажем, что $f \in \sigma\big(f^{-1}(\mathcal{K})\big)\big|\sigma(\mathcal{K})$. Для этого рассмотрим совокупность подмножеств множества Y, определенную равенством

$$
\mathcal{D} = \Big\{ D \subseteq Y \;\Big|\; f^{-1}(D) \in \sigma\big(f^{-1}(\mathcal{K})\big) \Big\}\;.
$$

Она является сигма-алгеброй подмножеств множества Y. Действительно, $Y \in \mathcal{D}$, так как $f^{-1}(Y) = X \in \sigma\big(f^{-1}(\mathcal{K})\big)$, так что выполнена аксиома $(\sigma - I)$. Для любых $A, B \in \mathcal{D}$ имеем $f^{-1}(A \setminus B) =$

$f^{-1}(A) \setminus f^{-1}(B) \in \sigma\big(f^{-1}(\mathcal{K})\big)$, поскольку $f^{-1}(A) \in \sigma\big(f^{-1}(\mathcal{K})\big)$ и $f^{-1}(B) \in \sigma\big(f^{-1}(\mathcal{K})\big)$, так что $A \setminus B \in \mathcal{D}$, а значит выполнена аксиома $(\sigma - II)$. Для любого набора $\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{D}$ имеем $f^{-1}\big(\bigcup_{i=1}^{\infty} A_i\big) = \bigcup_{i=1}^{\infty} f^{-1}(A_i) \in \sigma\big(f^{-1}(\mathcal{K})\big)$ в силу того, что $f^{-1}(A_i) \in \sigma\big(f^{-1}(\mathcal{K})\big)$ для любого i, так что $\bigcup_{i=1}^{\infty} A_i \in \mathcal{D}$. Таким образом, выполнена аксиома $(\sigma - III)$.

Сигма-алгебра $\mathcal{D}$ содержит $\mathcal{K}$, поэтому $\sigma(\mathcal{K}) \subseteq \mathcal{D}$ как минимальная сигма-алгебра, обладающая этим свойством. По определению $\mathcal{D}$ это означает, что

$$\forall\ B \in \sigma(\mathcal{K}) \ \ f^{-1}(B) \in \sigma\big(f^{-1}(\mathcal{K})\big)\,,$$

но так как по условию $f^{-1}(\mathcal{K}) \subseteq \mathcal{X}$, то $\sigma(\mathcal{K}) \subseteq \mathcal{X}$ как минимальная сигма-алгебра, содержащая систему $f^{-1}(\mathcal{K})$. Таким образом, доказано, что

$$\forall\ B \in \sigma(\mathcal{K}) \ \ f^{-1}(B) \in \mathcal{X}\,. \qquad \blacksquare$$

Следствие 5.1.3. Пусть $(X, \mathcal{X})$ — измеримое пространство. Функция $f : X \to \mathbb{R}$ измерима тогда и только тогда, когда выполнено условие

$$\forall\ c \in \mathbb{R} \ \ f^{-1}\big((-\infty; c)\big) \in \mathcal{X}\,.$$

Доказательство. Полагая $\mathcal{K} = \{(-\infty; c) \,|\, c \in \mathbb{R}\}$, пользуясь тем, что $\sigma(\mathcal{K}) = \mathcal{B}(\mathbb{R})$ и применяя предложение 5.1.2, получаем требуемое утверждение. $\blacksquare$

Отметим, что измеримость функций сохраняется при арифметических операциях над ними и при поточечном переходе к пределу.

5.2. Случайные величины и распределения вероятностей

Определение 5.2.1. Случайной величиной, определенной на вероятностном пространстве $(\Omega, \mathcal{F}, P)$, называется любое измеримое отображение $\xi : \Omega \to \mathbb{R}$, т. е. такое, что

$$\forall\ B \in \mathcal{B}(\mathbb{R}) \quad \xi^{-1}(B) = \{\omega \in \Omega \,|\, \xi(\omega) \in B\} \in \mathcal{F}\,. \tag{5.3}$$

Для случайной величины требование измеримости совершенно естественно. Ведь это требование того, чтобы любое множество исходов $\omega \in \Omega$ случайного эксперимента, при которых $\xi(\omega)$ попадает в то или иное борелевское подмножество $\mathbb{R}$ (в частности, в тот или иной промежуток), являлось событием, а значит для него была бы определена вероятность. Например, в случайном эксперименте с посадкой самолета предметом интереса могут быть такие события, как «вертикальная и горизонтальная составляющие скорости не превышают критических пороговых значений», «температура двигателя находится в промежутке, соответствующем оптимальному режиму работы» и т. п.

Из следствия 5.1.3 вытекает следующее

Предложение 5.2.1. Отображение $\xi : \Omega \to \mathbb{R}$ является случайной величиной тогда и только тогда, когда

$$\forall\ c \in \mathbb{R} \;\; \xi^{-1}\big((-\infty; c)\big) = \{\omega \in \Omega \,|\, \xi(\omega) < c\} \in \mathcal{F}\,. \tag{5.4}$$

В силу этого предложения условие (5.4) часто берут за определение случайной величины.

Из свойств измеримых функций следует, что результатом арифметических операций над случайными величинами и поточечного перехода к пределу в последовательностях случайных величин всегда является некоторая случайная величина.

R В дальнейшем случайные величины мы будем обозначать греческими буквами и, как это принято в теории вероятностей, опускать в записи аргумент ω. В частности, для обозначения события, состоящего в том, что случайная величина ξ приняла значение, принадлежащее B:

$$\{\omega \in \Omega \,|\, \xi(\omega) \in B\}\,, \quad B \in \mathcal{B}(\mathbb{R})$$

будем использовать более краткую запись: $\{\xi \in B\}$.

Определение 5.2.2. Пусть ξ — случайная величина, определенная на вероятностном пространстве $(\Omega, \mathcal{F}, P)$. ***Распределением вероятностей*** случайной величины ξ называют вероятностную меру P_ξ, определенную на борелевской σ-алгебре $\mathcal{B}(\mathbb{R})$ равенством

$$\forall\ B \in \mathcal{B}(\mathbb{R}) \quad P_\xi(B) = P(\xi \in B)\,. \tag{5.5}$$

Очевидно, что при этом тройка $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_\xi)$ является вероятностным пространством, т. е. удовлетворяет требованиям определения 2.1.1. Таким образом, каждая случайная величина порождает свое вероятностное пространство. В определенном смысле рассмотрение таких вероятностных пространств более удобно, ведь при этом в роли пространства элементарных событий Ω выступает не просто некая, подчас трудноописуемая совокупность исходов эксперимента (можно вспомнить тот же пример с посадкой самолета), а совокупность хорошо нам известных объектов — действительных чисел, а роль событий играют всевозможные промежутки и другие борелевские подмножества множества $\mathbb{R}$.

Оказывается, каждая вероятностная мера, определенная на $\mathcal{B}(\mathbb{R})$, является распределением вероятностей некоторой случайной величины. Справедливо следующее предложение.

Предложение 5.2.2. Пусть P_0 вероятностная мера, определенная на $\mathcal{B}(\mathbb{R})$. Тогда существуют вероятностное пространство $(\Omega, \mathcal{F}, P)$ и определенная на нем случайная величина ξ такие, что $P_\xi = P_0$.

Доказательство. Положим $\Omega = \mathbb{R}$, $\mathcal{F} = \mathcal{B}(\mathbb{R})$, $P = P_0$ и определим случайную величину ξ равенством $\xi(\omega) = \omega$ для любого $\omega \in \Omega = \mathbb{R}$. Тогда для любого $B \in \mathcal{B}(\mathbb{R})$ получим

$$P_\xi(B) = P(\xi \in B) = P_0\bigl(\{\omega \in \mathbb{R}\,|\,\omega \in B\}\bigr) = P_0(B)\,. \qquad \blacksquare$$

Таким образом, можно строить вероятностные модели, просто задавая распределения вероятностей на борелевской σ-алгебре $\mathcal{B}(\mathbb{R})$.

Далее мы рассмотрим наиболее часто встречающиеся в приложениях типы случайных величин и способы задания их распределений вероятностей.

5.3. Дискретные случайные величины

Дискретные случайные величины — простейший, но в то же время очень важный тип случайных величин. Они встречаются во многих приложениях, а кроме того, используются для аппроксимации более сложно устроенных случайных величин.

Определение 5.3.1. Дискретной случайной величиной, определенной на вероятностном пространстве $(\Omega, \mathcal{F}, P)$, называется случайная величина вида

$$\xi(\omega) = \sum_k x_k \mathbf{1}_{A_k}(\omega)\,, \tag{5.6}$$

где $\{x_k\} \subset \mathbb{R}$ — конечное или счетное множество, $\{A_k\} \subset \mathcal{F}$ — набор событий, символом $\mathbf{1}_A$ обозначен индикатор множества A. [a]

[a] Индикатором множества A называется функция, определенная на Ω равенством

$$\mathbf{1}_A(\omega) = \begin{cases} 1\,, & \omega \in A\,, \\ 0\,, & \omega \notin A\,. \end{cases}$$

Нетрудно заметить, что формула (5.6) при указанных в определении условиях всегда задает случайную величину, т. е. измеримое отображение из Ω в $\mathbb{R}$. Действительно, для любого $B \in \mathcal{B}(\mathbb{R})$ имеем:

$$\xi^{-1}(B) = \{\omega \in \Omega \,|\, \xi(\omega) \in B\} = \bigcup_{x_k \in B} A_k \in \mathcal{F}\,,$$

поскольку все A_k принадлежат сигма-алгебре $\mathcal{F}$.

R Заметим, что представление дискретной случайной величины в виде (5.6) неоднозначно. Например, рассмотрим случайный эксперимент, заключающийся в бросании одной игральной кости. Пусть

$$A = \{\text{выпало четное число}\}\,,$$
$$A_i = \{\text{выпало число } i\}\,, \quad i = 1, 2, \ldots, 6\,.$$

Тогда

$$\xi(\omega) := \mathbf{1}_A(\omega) = 1 \cdot \mathbf{1}_A(\omega) + 0 \cdot \mathbf{1}_{\overline{A}}(\omega) = \sum_{k=1}^{6} x_k \mathbf{1}_{A_k}(\omega)\,,$$

где $x_k = 1$, если k четное, и $x_k = 0$, если k нечетное.

Представление (5.6) становится однозначным, если потребовать выполннение условий:

- значения x_k попарно различны;
- события $\{A_k\}$ образуют полную группу событий.

В таком случае совокупность всех x_k образует множество значений случайной величины ξ, а A_k представляет собой событие, состоящее в том, что ξ приняла значение x_k:

$$A_k = \{\xi = x_k\} .$$

Если не оговорено иное, в дальнейшем всегда будем предполагать эти требования выполненными.

Для дискретной случайной величины вида (5.6) введем обозначения:

$$p_k = P(\xi = x_k) .$$

Очевидно, эти вероятности удовлетворяют соотношению

$$\sum_k p_k = 1 . \tag{5.7}$$

Предложение 5.3.1. Значения p_k полностью определяют распределение вероятностей дискретной случайной величины ξ, а именно для любого $B \in \mathcal{B}(\mathbb{R})$

$$P_\xi(B) = \sum_{x_k \in B} p_k . \tag{5.8}$$

Доказательство. Действительно, имеем:

$$P_\xi(B) = P(\xi \in B) = P\left(\bigsqcup_{x_k \in B} \{\xi = x_k\}\right) =$$

$$= \sum_{x_k \in B} P\left(\{\xi = x_k\}\right) = \sum_{x_k \in B} p_k . \qquad \blacksquare$$

Рассмотрим некоторые часто используемые дискретные распределения вероятностей.

▪ Пример 5.3.1. — Дискретное равномерное распределение.

Обозначение: $\boxed{\xi \sim U\{x_1, x_2, \ldots, x_n\}}$

Определение 5.3.2. Случайная величина ξ называется ***равномерно распределенной по значениям $\boldsymbol{x_1, x_2, \ldots, x_n}$***, если в результате проведения случайного эксперимента она может принять любое из этих значений с вероятностью $p_k = \dfrac{1}{n}$, $k = 1, 2, \ldots, n$, $n \in \mathbb{N}$.

Например, пусть ξ — число очков, полученное при броске игральной кости. Тогда $\xi \sim U\{1, 2, 3, 4, 5, 6\}$. ▪

▪ Пример 5.3.2. — Гипергеометрическое распределение вероятностей.

Обозначение: $\boxed{\xi \sim HG(M, N, n)}$

Определение 5.3.3. Случайная величина ξ имеет ***гипергеометрическое распределение вероятностей с параметрами $\boldsymbol{M, N, n}$***, где $M, N, n \in \mathbb{N}$, $0 < M < N$, $0 < n < N$, если она принимает значения $k \in \mathbb{Z}$, где $\max\{0; n - (N - M)\} \leq k \leq \min\{n; M\}$ с вероятностями

$$p_k = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}. \tag{5.9}$$

Гипергеометрическое распределение возникает в следующем случайном эксперименте. Пусть в урне находится N шаров, из которых M — черные, а $N - M$ — белые. Из урны наудачу извлекают n шаров. Пусть ξ — количество черных шаров среди извлеченных. Тогда $\xi \sim HG(M, N, n)$. Действительно, минимальное значение, которое может принять ξ, равно 0, если количество извлекаемых шаров n меньше либо равно $N - M$ — количества белых шаров. Если же $n > N - M$, минимальное возможное количество черных шаров среди извлеченных получается, если извлечены все имеющиеся белые, и равно $n - (N - M)$. Максимальное значение, которое может принять ξ, равно n, если $n < M$, и равно M — количеству черных шаров, если $n \geq M$. Равенство (5.9) для вероятностей p_k можно получить, заметив, что данный случайный эксперимент описывается схемой с конечным числом равновероятных исходов. Исход эксперимента можно отождествить с сочетанием из N имеющихся в урне шаров по n, поэтому общее количество исходов равно C_N^n. При этом число сочетаний, содержащих ровно k черных шаров, вычисляется по правилу произведения: оно равно произведению C_M^k (числа способов, которыми можно извлечь k шаров из имеющихся M черных) на C_{N-M}^{n-k} (число способов, которыми можно выбрать оставшиеся $n-k$ шаров из имеющихся в урне $N-M$ белых). ■

■ Пример 5.3.3. — Биномиальное распределение вероятностей.

Обозначение: $\boxed{\xi \sim B(n,p)}$

Определение 5.3.4. Случайная величина ξ имеет ***биномиальное распределение вероятностей с параметрами n и p***, где $n \in \mathbb{N}$, $p \in (0;1)$, если она принимает значения $k \in \{0,1,\ldots,n\}$ с вероятностями

$$p_k = C_n^k p^k q^{n-k}\,, \text{ где } q = 1-p\,. \tag{5.10}$$

Биномиальное распределение имеет число успехов в схеме Бернулли, где n — число испытаний, а p — вероятность успеха. ■

R Биномиальное распределение используют для аппроксимации гипергеометрического при больших значениях параметров N и M. Это возможно, так как справедливо следующее предложение.

Предложение 5.3.2. При фиксированном значении параметра n, если $N, M \to \infty$ так, что $M/N \to p$, гипергеометрическое распределение вероятностей с параметрами M, N, n сходится к биномиальному с параметрами n, p, т. е.

$$\lim_{\substack{N,M\to\infty \\ \frac{M}{N}\to p}} \frac{C_M^k C_{N-M}^{n-k}}{C_N^n} = C_n^k p^k q^{n-k}\,.$$

Доказательство. Преобразуем дробь и сократим общие множители факториалов:

$$\frac{C_M^k C_{N-M}^{n-k}}{C_N^n} = \frac{M!\,(N-M)!\,n!\,(N-n)!}{k!\,(M-k)!\,(n-k)!\,(N-M-n+k)!\,N!} =$$
$$= \frac{n!}{k!\,(n-k)!} \cdot \frac{M!\,(N-M)!\,(N-n)!}{(M-k)!\,(N-M-n+k)!\,N!} =$$
$$= C_n^k \cdot \frac{\overbrace{M\ldots(M-k+1)}^{k}\overbrace{(N-M)\ldots(N-M-n+k+1)}^{n-k}}{\underbrace{N(N-1)\ldots(N-n+1)}_{n}}\,.$$

Разделив каждый множитель числителя на каждый множитель знаменателя, получаем

$$\frac{C_M^k C_{N-M}^{n-k}}{C_N^n} = C_n^k \cdot \frac{M}{N} \cdot \frac{M-1}{N-1} \cdot \frac{M-k+1}{N-k+1} \cdot \frac{N-M}{N-k} \cdot$$
$$\cdot \frac{N-M-1}{N-k-1} \cdot \ldots \cdot \frac{N-M-n+k+1}{N-n+1}\,. \tag{5.11}$$

Для любых $C, L \in \mathbb{R}$ имеем:

$$\lim_{\substack{N,M\to\infty\\ \frac{M}{N}\to p}} \frac{M-C}{N-C} = \lim_{\substack{N,M\to\infty\\ \frac{M}{N}\to p}} \frac{\frac{M}{N}-\frac{C}{N}}{1-\frac{C}{N}} = p\,,$$
$$\lim_{\substack{N,M\to\infty\\ \frac{M}{N}\to p}} \frac{N-M-C}{N-L} = \lim_{\substack{N,M\to\infty\\ \frac{M}{N}\to p}} \frac{1-\frac{M}{N}-\frac{C}{N}}{1-\frac{L}{N}} = 1-p\,.$$

Поскольку k и n фиксированы, каждый из первых k множителей в правой части равенства (5.11) при $N \to \infty$ и $\dfrac{M}{N} \to p$ дает в пределе p, а каждый из оставшихся $n-k$ множителей дает в пределе $1-p$. ∎

Пусть, например, из совокупности всех жителей большого города выбирается n человек, случайная величина ξ принимает значение, равное числу женщин в выборке. Очевидно, $\xi \sim HG(M, N, n)$, где N — количество жителей города, M — количество женщин среди жителей города. Однако, поскольку N и M велики, можно считать, что $\xi \sim B(n, p)$, где $p = M/N$ — доля женщин среди горожан (см. задачу 7 из раздела 4.3).

▪ Пример 5.3.4. — распределение Пуассона.

Обозначение: $\boxed{\xi \sim P(\lambda)}$

Определение 5.3.5. Случайная величина ξ ***распределена по закону Пуассона с параметром $\boldsymbol{\lambda}$***, если она принимает значения $k \in \{0\} \cup \mathbb{N}$ с вероятностями

$$p_k = \frac{\lambda^k}{k!} e^{-\lambda}. \tag{5.12}$$

Распределение Пуассона является предельным по отношению к биномиальному. Это утверждает доказанная в предыдущей главе теорема Пуассона (теорема 4.2.1). Оно используется для построения математических моделей потоков случайных событий. Например, рассмотрим математическую модель радиоактивного распада.

Пусть имеется образец радиоактивного материала, испускающий в случайные моменты времени радиоактивные частицы, появляющиеся в результате радиоактивного распада атомов этого образца. Пусть ξ — количество частиц, испускаемых образцом за единицу времени. Предположим, что N — количество атомов, из которых состоит образец. Будем считать, что каждый атом за единицу времени распадается независимо от других с вероятностью p. Наблюдение за радиоактивным образцом в течение единичного промежутка времени при этих условиях представляет собой схему Бернулли, где отдельным испытанием является наблюдение за одним из атомов, успехом — его распад, так что ξ — количество успехов в серии из N испытаний Бернулли с вероятностью успеха p. Следовательно, эта случайная величина должна быть распределена по закону Бернулли с параметрами N, p. Поскольку N в данном случае очень велико, а p чрезвычайно мало, в силу теоремы Пуассона, естественно считать, что $\xi \sim P(\lambda)$, где $\lambda = pN$, что можно интерпретировать как среднее количество частиц, испускаемых за единицу времени. ∎

▪ Пример 5.3.5. — Геометрическое распределение вероятностей.

Обозначение: $\boxed{\xi \sim \mathrm{Geom}(p)}$

Определение 5.3.6. Случайная величина ξ имеет ***геометрическое распределение вероятностей с параметром $\boldsymbol{p}$***, где $p \in (0; 1)$, если она принимает значения $k \in \mathbb{N}$ с вероятностями

$$p_k = q^{k-1}p\,, \text{ где } q = 1 - p\,. \tag{5.13}$$

Рассмотрим серию одинаковых независимых испытаний с вероятностью успеха p, которая проводится до наступления первого успеха. Пусть ξ — номер испытания, при котором произошел первый успех (количество испытаний в такой серии). Тогда $\xi \sim \mathrm{Geom}(p)$. Действительно, обозначим

$$A_i = \{i\text{-е испытание закончилось успехом}\}\,, \quad i \in \mathbb{N}\,,$$
$$B_k = \{\text{первый успех произошел при } k\text{-м испытании}\}\,, \quad k \in \mathbb{N}\,.$$

Тогда $B_k = \overline{A}_1 \ldots \overline{A}_{k-1} A_k$ и, в силу независимости событий A_i, имеем:

$$P(B_k) = P(\overline{A}_1) \ldots P(\overline{A}_{k-1}) P(A_k) = q^{k-1}p\,.$$

■

5.4. Функция распределения случайной величины

Рассмотрим вероятностное пространство, построенное в задаче о встрече (Задача 1.5.1, см. также пример 2.1.2). Пусть ξ — момент прихода на место встречи первого участника эксперимента. Эта случайная величина не является дискретной, так как может принимать любое значение из промежутка $[0; 1]$. При этом для любого $x \in [0; 1]$ вероятность события $\{\xi = x\}$ равна нулю. Поэтому описание распределения вероятностей такой случайной величины через вероятности ее значений невозможно. Рассмотрим другой подход к описанию распределений вероятностей, опирающийся на понятие ***функции распределения***.

Определение 5.4.1. Пусть ξ — случайная величина, определенная на некотором вероятностном пространстве $(\Omega, \mathcal{F}, P)$. ***Функция распределения вероятностей*** ξ обозначается F_ξ и определяется равенством:

$$\forall\, x \in \mathbb{R} \quad F_\xi(x) = P(\xi < x)\,. \tag{5.14}$$

Определение функции распределения корректно, так как по определению случайной величины для любого $x \in \mathbb{R}$ совокупность исходов $\{\omega \,|\, \xi(\omega) \in (-\infty; x)\}$ является событием, элементом сигма-алгебры $\mathcal{F}$, а значит, для него определена вероятность.

▪ Пример 5.4.1. Пусть $\xi \sim U\{1; 2; 3\}$. Тогда (см. рис. 5.1)

$$F_\xi(x) = \begin{cases} 0 & x \leq 1\,, \\ 1/3\,, & 1 < x \leq 2\,, \\ 2/3\,, & 2 < x \leq 3\,, \\ 1\,, & x > 3\,. \end{cases}$$ ▪

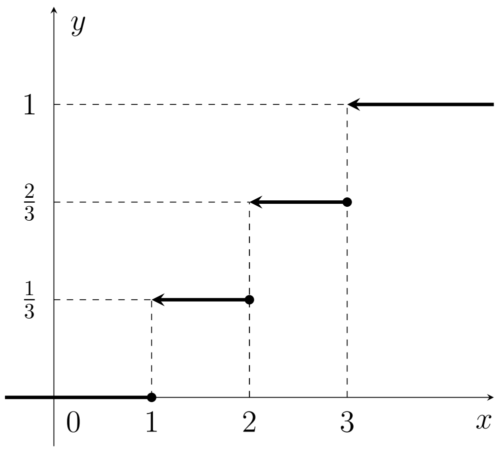


Рис. 5.1. График функции распределения случайной величины $\xi \sim U\{1; 2; 3\}$

▪ Пример 5.4.2. Рассмотрим вероятностное пространство $(\Omega, \mathcal{F}, P)$, построенное в задаче о встрече ($\Omega = [0; 1] \times [0; 1]$, $\mathcal{F}$ — сигма-алгебра измеримых по Лебегу подмножеств множества Ω, P — мера Лебега на плоскости). Пусть ξ — момент прихода первого участника эксперимента, а η — момент прихода второго, т. е. $\xi(\omega) = u$, $\eta(\omega) = v$, где $\omega = (u; v) \in \Omega$. Тогда (см. рис. 5.2)

$$F_\xi(x) = F_\eta(x) = \begin{cases} 0\,, & x \leq 0\,, \\ x\,, & 0 < x \leq 1\,, \\ 1\,, & x > 1\,. \end{cases}$$

▪

Рассмотрим свойства функций распределения.

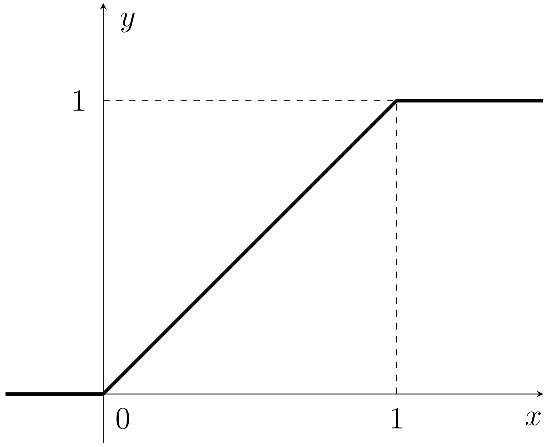


Рис. 5.2. График функции распределения момента прихода участника в задаче о встрече

Предложение 5.4.1. Пусть $F(x)$ — функция распределения некоторой случайной величины. Тогда

(F-I) $\forall\ x \in \mathbb{R}\ \ 0 \leq F(x) \leq 1$;

(F-II) $F(x)$ возрастает (нестрого);

(F-III) $F(-\infty) = 0\,, F(+\infty) = 1$;

(F-IV) $\forall\ x \in \mathbb{R}\ \ F(x-0) = F(x)$, т. е. $F(x)$ непрерывна слева в каждой точке $\mathbb{R}$.

Доказательство. Свойство (F-I) очевидно. Свойство (F-II) следует из свойства монотонности вероятности (свойство (P-4)):

$$x_1 < x_2 \ \Rightarrow\ \{\xi < x_1\} \subseteq \{\xi < x_2\} \ \Rightarrow\ F_\xi(x_1) = \\ = P(\xi < x_1) \leq P(\xi < x_2) = F_\xi(x_2)\,.$$

Для доказательства свойства (F-III) возьмем $\{x_n\} \subset \mathbb{R}$ так, что $\lim\limits_{n\to\infty} x_n = +\infty$. Без ограничения общности будем считать, что x_n возрастает. Тогда по свойству непрерывности вероятности (P-9) для любого $n \in \mathbb{N}$ имеем $\{\xi < x_n\} \subseteq \{\xi < x_{n+1}\}$, следовательно

$$\lim_{n\to\infty} F_\xi(x_n) = \lim_{n\to\infty} P(\xi < x_n) = P\left(\sum_{n=1}^{\infty}\{\xi < x_n\}\right) = \\ = P\left(\xi \in \bigcup_{n=1}^{\infty}(-\infty; x_n)\right) = P(\xi \in \mathbb{R}) = P(\Omega) = 1\,.$$

Если $\lim\limits_{n\to\infty} x_n = -\infty$, предполагая без ограничения общности, что $\{x_n\}$ убывает, используя то, что при этом $\{\xi < x_n\} \supseteq \{\xi < x_{n+1}\}$ для любого $n \in \mathbb{N}$, по второму утверждению свойства (P-9) получаем:

$$\lim_{n\to\infty} F_\xi(x_n) = \lim_{n\to\infty} P(\xi < x_n) = P\left(\prod_{n=1}^{\infty}\{\xi < x_n\}\right) =$$
$$= P\left(\xi \in \bigcap_{n=1}^{\infty}(-\infty; x_n)\right) = P(\xi \in \varnothing) = P(\varnothing) = 0\,.$$

Докажем свойство (F-IV). Пусть $\{x_n\} \subset \mathbb{R}$, $\lim\limits_{n\to\infty} x_n = x - 0$. Без ограничения общности будем считать, что x_n возрастает. Тогда по свойству (P-9), из того, что $\{\xi < x_n\} \subseteq \{\xi < x_{n+1}\}$ для любого $n \in \mathbb{N}$, следует

$$\lim_{n\to\infty} F_\xi(x_n) = \lim_{n\to\infty} P(\xi < x_n) = P\left(\sum_{n=1}^{\infty}\{\xi < x_n\}\right) =$$
$$= P\left(\xi \in \bigcup_{n=1}^{\infty}(-\infty; x_n)\right) = P(\xi \in (-\infty; x)) = F_\xi(x)\,. \quad \blacksquare$$

R Функция распределения случайной величины не обязательно непрерывна справа всюду в $\mathbb{R}$. В качестве контрпримера можно взять функцию распределения дискретной случайной величины (см. пример 5.4.1).

Справедлива теорема, в некотором смысле, обратная к доказанному предложению.

Теорема 5.4.2. Пусть функция $F : \mathbb{R} \to \mathbb{R}$ удовлетворяет условиям (F-I), (F-II), (F-III) и (F-IV). Тогда существует вероятностное пространство $(\Omega, \mathcal{F}, P)$ и определенная на нем случайная величина ξ такие, что $F_\xi(x) = F(x)$.

Доказательство. Возьмем $\Omega = [0; 1]$. Пусть $\mathcal{F}$ — сигма-алгебра измеримых по Лебегу подмножеств промежутка $[0; 1]$, $P = \mu_L$ — мера Лебега в $\mathbb{R}$.

1. Предположим, что $F(x)$ строго возрастает. Тогда существует обратная функция $F^{-1}(x)$. Положим

$$\forall\, \omega \in \Omega \;\; \xi(\omega) = F^{-1}(\omega)\,,$$

тогда для любого $x \in \mathbb{R}$

$$F_\xi(x) = P\big(\{\omega \in \Omega \,|\, \xi(\omega) < x\}\big) = P\big(\{\omega \in \Omega \,|\, F^{-1}(\omega) < x\}\big) = \\ = P\big(\{\omega \in \Omega \,|\, \omega < F(x)\}\big) = \mu_L\big([0; F(x))\big) = F(x)\,.$$

2. Если $F(x)$ не является строго возрастающей, она необратима. В этом случае доказательство проводится аналогично, только вместо обратной к F используется обобщенная обратная функция F^{inv}, которая определяется равенством

$$F^{inv}(y) = \inf\{x \in \mathbb{R} \,|\, F(x) > y\}\,, \quad y \in [0;1]\,. \qquad \blacksquare$$

Функция распределения F_ξ описывает распределение вероятностей случайной величины ξ по лучам $(-\infty; x)$, где $x \in \mathbb{R}$ (значение $F_\xi(x)$ — вероятность попадания ξ в луч $(-\infty; x)$). Оказывается, задание этих вероятностей полностью определяет P_ξ — распределение вероятностей ξ по всем борелевским подмножествам $\mathbb{R}$. Этот нетривиальный факт, который мы принимаем без доказательства (доказательство можно найти, например в [8]), опирается, в частности, на то, что сигма-алгебра борелевских подмножеств $\mathbb{R}$ — минимальная сигма-алгебра, содержащая все лучи $(-\infty; x)$, где $x \in \mathbb{R}$ (см. предложение 5.1.1). Для практической работы с функцией распределения важен более тривиальный факт, состоящий в том, что знание функции распределения случайной величины позволяет находить вероятности попадания ее значений в любые промежутки числовой прямой, а именно справедливы следующие формулы:

$$P(\xi \geq a) = 1 - F_\xi(a)\,; \tag{5.15}$$
$$P(\xi \leq b) = F_\xi(b+0)\,; \tag{5.16}$$
$$P(\xi > a) = 1 - F_\xi(a+0)\,; \tag{5.17}$$
$$P(a \leq \xi < b) = F_\xi(b) - F_\xi(a)\,; \tag{5.18}$$
$$P(a < \xi < b) = F_\xi(b) - F_\xi(a+0)\,; \tag{5.19}$$
$$P(a \leq \xi \leq b) = F_\xi(b+0) - F_\xi(a)\,; \tag{5.20}$$
$$P(a < \xi \leq b) = F_\xi(b+0) - F_\xi(a+0)\,. \tag{5.21}$$

Формула (5.15) доказывается с помощью свойства вероятности противоположного события (Р-6):

$$P(\xi \geq a) = P\left(\overline{\{\xi < a\}}\right) = 1 - P\left(\xi < a\right) = 1 - F_\xi(a)\,.$$

Для доказательства формулы (5.16) используем свойство непрерывности вероятности (P-9). В силу вложения

$$\left\{\xi < b + \frac{1}{n+1}\right\} \subseteq \left\{\xi < b + \frac{1}{n}\right\},$$

имеем:

$$P(\xi \le b) = P\left(\xi \in \bigcap_{n\in\mathbb{N}} \left(-\infty; b + \frac{1}{n}\right)\right) = P\left(\prod_{n\in\mathbb{N}} \left\{\xi < b + \frac{1}{n}\right\}\right) =$$
$$= \lim_{n\to\infty} P\left(\xi < b + \frac{1}{n}\right) = \lim_{n\to\infty} F_\xi\left(b + \frac{1}{n}\right) = F_\xi(b+0)\,.$$

Формула (5.17) доказывается аналогично (5.15) с использованием формулы (5.16).

Формула (5.18) следует из свойства (P-3), т. к. $\{\xi < a\} \subseteq \{\xi < b\}$ при $a < b$. Формулы (5.19), (5.20) и (5.21) доказываются аналогично, с использованием формулы (5.16).

R В литературе по теории вероятностей можно встретить альтернативное определение функции распределения: $F_\xi(x) = P(\xi \le x)$. При таком определении для функции распределения остаются справедливыми свойства (F-I), (F-II) и (F-III), а вместо свойства (F-IV) будет иметь место непрерывность справа во всех точках $\mathbb{R}$. Аналоги формул (5.15)—(5.21) будут отличаться тем, что везде в них правосторонние пределы заменятся на значения F_ξ в соответствующих точках, а значения F_ξ заменятся на левосторонние пределы в соответствующих точках, например, вместо формулы (5.19) будет верно равенство $P(a < \xi < b) = F_\xi(b-0) - F_\xi(a)$.

Определение 5.4.2. ***Квантилем порядка*** $\boldsymbol{\alpha}$ распределения вероятностей случайной величины ξ называется такое число $z_\alpha \in \mathbb{R}$, что выполнены условия

$$P(\xi \le z_\alpha) \ge \alpha\,, \quad P(\xi \ge z_\alpha) \ge 1-\alpha\,.$$

Квантили порядка $0,25$, $0,5$ и $0,75$ называются соответственно первым, вторым и третьим ***квартилями*** распределения вероятностей случайной величины ξ. Второй квартиль называется ***медианой***. Квантили порядка $0,1$, $0,2$, $\ldots$, $0,9$ называются ***децилями***. Квантили порядка $0,01$, $0,02$, $0,03, \ldots, 0,99$ называются

процентилями.

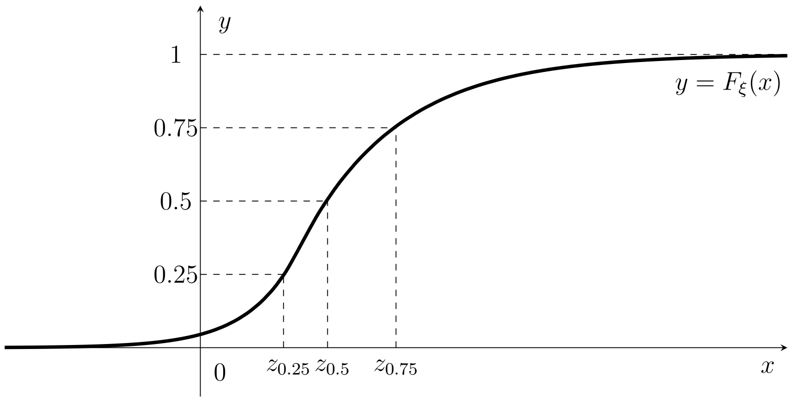


Рис. 5.3. Квартили распределения вероятностей случайной величины

Квантиль порядка α распределения вероятностей случайной величины ξ определяется однозначно, если функция распределения $F_\xi(x)$ непрерывна. В этом случае справедливо равенство

$$z_\alpha = F_\xi^{-1}(\alpha)\,.$$

Взаимное расположение квантилей одного порядка на числовой прямой является одной из удобных наглядных характеристик распределения вероятностей (см. рис. 5.3.).

5.5. Абсолютно непрерывные случайные величины

Рассмотрим еще один способ описания распределения вероятностей случайной величины.

Определение 5.5.1. Случайная величина ξ называется ***абсолютно непрерывной***, если существует такая неотрицательная интегрируемая функция $f_\xi : \mathbb{R} \to \mathbb{R}$, что для любого $x \in \mathbb{R}$ выполнено равенство

$$F_\xi(x) = \int\limits_{-\infty}^{x} f_\xi(t)\,dt\,. \tag{5.22}$$

При этом функция f_ξ называется ***плотностью распределения вероятностей*** случайной величины ξ.

Предложение 5.5.1. Если ξ — абсолютно непрерывная случайная величина, то ее функция распределения F_ξ непрерывна во всех точках $\mathbb{R}$, а ее плотность распределения f_ξ обладает свойствами

(f-I) $\forall\ x \in \mathbb{R}\ \ f(x) \geq 0\,;$

(f-II) $\displaystyle\int\limits_{-\infty}^{+\infty} f_\xi(x)\,dx = 1\,;$

(f-III) В точках непрерывности функции f_ξ функция распределения F_ξ дифференцируема, и справедливо равенство: $F'_\xi(x) = f_\xi(x)\,;$

(f-IV) $\forall\ a, b \in \mathbb{R}\ (a \leq b):\quad P(a \leq \xi \leq b) = P(a < \xi \leq b) =$

$$= P(a \leq \xi < b) = P(a < \xi < b) = \int\limits_a^b f_\xi(x)\,dx\,.$$

Доказательство. Свойство (f-I) заложено в определении 5.5.1. Свойство (f-II) следует из равенства (5.22) и свойства (F-III) функции распределения. Непрерывность F_ξ и свойство (f-III) вытекают из известных из математического анализа свойств интеграла с переменным верхним пределом. Свойство (f-IV) получается из формул (5.15)—(5.21) благодаря непрерывности F_ξ и свойству (f-III) с помощью формулы Ньютона–Лейбница. ■

Предложение 6.3.1 проясняет смысл термина «плотность распределения вероятностей». График функции f_ξ (см. рис. 5.4.) показывает как единичная вероятностная масса «размазана» по числовой прямой: площадь под графиком равна 1, при этом площадь криволинейной трапеции, приходящаяся на промежуток с концами a и b (отрезок, интервал, или полуинтервал), равна вероятности попадания ξ в этот промежуток.

Рассмотрим некоторые часто встречающиеся в приложениях распределения вероятностей абсолютно непрерывных случайных величин.

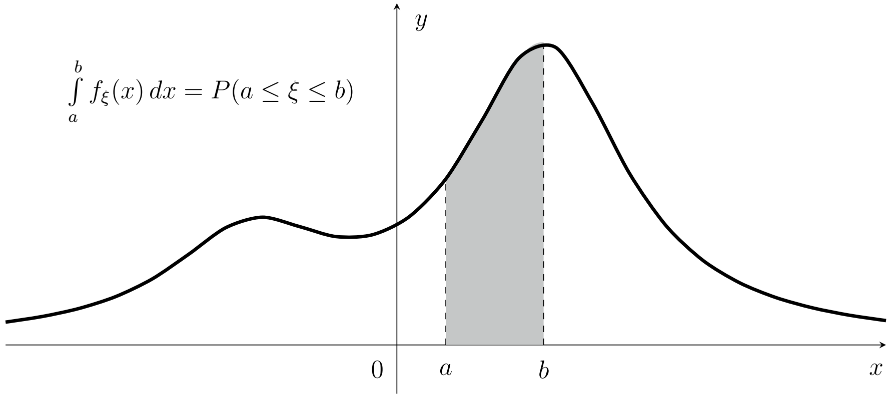


Рис. 5.4. Плотность распределения вероятностей случайной величины

■ Пример 5.5.1. — Равномерное распределение вероятностей.

Обозначение: $\boxed{\xi \sim U(a, b)}$

Определение 5.5.2. Случайная величина ξ называется ***равномерно распределенной на отрезке*** $[\boldsymbol{a}; \boldsymbol{b}]$, если ее плотность распределения имеет вид

$$f_\xi(x) = \frac{1}{b-a} \cdot \mathbf{1}_{[a;b]}(x)\,.$$

Нетрудно проверить, что функция распределения при этом имеет вид

$$F_\xi(x) = \begin{cases} 0\,, & x < a\,, \\ \dfrac{x-a}{b-a}\,, & a \le x < b\,, \\ 1\,, & x \ge b\,. \end{cases}$$

В примере 5.4.2 было показано, что если ξ и η — моменты прихода участников эксперимента в задаче о встрече, то $\xi, \eta \sim U(0;1)$. ■

■ Пример 5.5.2. — Показательное распределение вероятностей.

Обозначение: $\boxed{\xi \sim \mathrm{Exp}(\lambda)}$

Определение 5.5.3. Случайная величина ξ имеет ***показательное распределение с параметром*** $\boldsymbol{\lambda}$, если

$$f_\xi(x) = \lambda e^{-\lambda x} \cdot \mathbf{1}_{[0;+\infty)}(x)\,.$$

Функция распределения ξ при этом имеет вид

$$F_\xi(x) = (1 - e^{-\lambda x}) \cdot \mathbf{1}_{[0;+\infty)}(x)\,.$$

Показательное распределение используется для моделирования ***времени ожидания при отсутствии последействия***. Пусть ξ — время исправной работы некоторого прибора (например, микросхемы) в ситуации, когда его выход из строя не связан с такими явлениями, как износ или старение, а может быть вызван только случайными внешними воздействиями (в случае микросхемы перепадом напряжения в цепи). Пусть $g(x) = P(\xi \geq x)$, т. е. вероятность того, что прибор оставался исправным до момента x. Эту функцию называют функцией надежности. Из сделанного предположения следует равенство

$$P(\xi \geq x + t | \xi \geq t) = P(\xi \geq x)\,, \quad x, t > 0\,, \tag{5.23}$$

т. е. если до момента t прибор исправен, то вероятность того, что он останется исправным до момента $x + t$, совпадает с вероятностью того, что за время x выхода из строя не произойдет. Это и называют ***отсутствием последействия***. Пользуясь равенством (5.23), получаем для функции g:

$$\begin{aligned} g(x) = P(\xi \geq x) = P(\xi \geq x + t | \xi \geq t) = \\ = \frac{P\big(\{\xi \geq x + t\} \cdot \{\xi \geq t\}\big)}{P(\xi \geq t)} = \frac{P(\xi \geq x + t)}{P(\xi \geq t)} = \frac{g(x+t)}{g(t)}\,. \end{aligned}$$

Таким образом, для любых $x, t > 0$ получаем

$$g(x + t) = g(x)g(t)\,.$$

Из этого равенства следуют соотношения

$$\forall\, n \in \mathbb{N} \quad g(n) = g(1)^n\,, \quad g\left(\frac{1}{n}\right)^n = g(1)\,,$$

и, как следствие,

$$\forall\, m, n \in \mathbb{N} \quad g\left(\frac{m}{n}\right) = g(1)^{\frac{m}{n}}\,.$$

Пусть $g(1) = e^{-\lambda}$, где $\lambda > 0$. Такое значение λ существует, поскольку $0 < g(1) < 1$. В результате $g\left(\frac{m}{n}\right) = e^{-\lambda\frac{m}{n}}$. Доопределяя функцию $g(x)$ по непрерывности в иррациональных точках промежутка $[0; +\infty)$, получаем $g(x) = e^{-\lambda x}$ при $x \geq 0$. Таким образом, при $x > 0$ получаем

$$F_\xi(x) = P(\xi < x) = 1 - P(\xi \geq x) = 1 - g(x) = 1 - e^{-\lambda x}\,,$$

следовательно, $f_\xi(x) = \dfrac{d}{dx}F_\xi(x) = \lambda e^{-\lambda x}$. ■

Предложение 5.5.2. Пусть ξ — абсолютно непрерывная случайная величина, а $g : \mathbb{R} \to \mathbb{R}$ — монотонная дифференцируемая функция. Тогда $\eta = g(\xi)$ — абсолютно непрерывная случайная величина, для плотности распределения вероятностей которой справедливо равенство

$$f_\eta(x) = f_\xi\big(g^{-1}(x)\big)\frac{d}{dx}g^{-1}(x)\,. \tag{5.24}$$

Доказательство. Пусть, без ограничения общности, функция g возрастает на $\mathbb{R}$, тогда для любого $x \in \mathbb{R}$

$$F_\eta(x) = P(\eta < x) = P(g(\xi) < x) = P(\xi < g^{-1}(x)) = F_\xi\big(g^{-1}(x)\big)\,.$$

Дифференцируя обе части этого равенства с помощью формулы производной сложной функции в точках множества, являющегося образом при отображении g множества точек непрерывности функции f_ξ, получаем равенство (5.24). ■

■ Пример 5.5.3. Пусть $\xi \sim U(0, 1)$, $\eta = -\dfrac{1}{\lambda}\ln(1 - \xi)$. Тогда $\eta \sim Exp(\lambda)$. Действительно, пусть $g(x) = -\dfrac{1}{\lambda}\ln(1 - x)$. Тогда $g^{-1}(x) = 1 - e^{-\lambda x}$ и по формуле (5.24) получаем

$$f_\eta(x) = f_\xi\big(g^{-1}(x)\big)\frac{d}{dx}g^{-1}(x) = \mathbf{1}_{[0;1]}\big(1 - e^{-\lambda x}\big)\lambda e^{-\lambda x} = \lambda e^{-\lambda x}\mathbf{1}_{[0;+\infty)}(x)\,,$$

поскольку $1 - e^{-\lambda x} \in [0; 1] \;\Leftrightarrow\; x \in [0; +\infty)$. ■

■ Пример 5.5.4. — Нормальное (гауссовское) распределение вероятностей.

Обозначение: $\boxed{\xi \sim N(a, \sigma^2)}$

Определение 5.5.4. Случайная величина ξ ***распределена по нормальному закону с параметрами $\boldsymbol{a}$, $\boldsymbol{\sigma^2}$***, если ее плотность распределения имеет вид

$$f_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}\,.$$

Распределение вероятностей случайной величины $\xi \sim N(0,1)$ называется ***стандартным нормальным***. Плотность стандартного нормального распределения вероятностей — функция Гаусса:

$$f_\xi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,.$$

Нормальное распределение вероятностей появляется как предельное по отношению к биномиальному в теоремах Муавра–Лапласа (теоремы 4.2.2 и 4.2.3). Они утверждают, что распределение вероятностей случайной величины $\dfrac{\xi - np}{\sqrt{npq}}$, где ξ — число успехов в серии из n испытаний Бернулли с вероятностью успеха p, при фиксированном p и $n \to \infty$ стремится к стандартному нормальному.

Гауссовские случайные величины используются для описания случайных ошибок, возникающих при различных измерениях, отклонений от цели при срельбах и других подобных явлений. Приведем некоторые их важные свойства.

Следующее предложение описывает свойство нормально распределенных случайных величин, связанное с линейными операциями над ними.

Предложение 5.5.3.

$$\xi \sim N(a, \sigma^2) \ \Leftrightarrow \ \xi = \sigma\eta + a\,, \text{ где } \eta \sim N(0,1)\,.$$

Доказательство. Достаточность. Пусть $\eta \sim N(0,1)$, $\xi = \sigma\eta + a$. Тогда, пользуясь предложением 5.5.2 и полагая в формуле (5.24)

$g(x) = \sigma x + a$, а, следовательно, $g^{-1}(x) = \dfrac{x-a}{\sigma}$, получаем

$$f_\xi(x) = f_\eta\big(g^{-1}(x)\big)\frac{d}{dx}g^{-1}(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}\,.$$

Необходимость. Пусть $\xi \sim N(a, \sigma^2)$. Положим $\eta = \dfrac{\xi - a}{\sigma}$. Тогда, снова пользуясь предложением 5.5.2, но теперь полагая в формуле (5.24) $g(x) = \dfrac{x-a}{\sigma}$, а, следовательно, $g^{-1}(x) = \sigma x + a$, получаем

$$f_\eta(x) = f_\xi\big(g^{-1}(x)\big)\frac{d}{dx}g^{-1}(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\,,$$

т. е. $\eta \sim N(0,1)$. При этом $\xi = \sigma\eta + a$. ■

Из предложения 5.5.3 вытекает

Следствие 5.5.4. Линейная функция от гауссовской случайной величины — гауссовская случайная величина.

▪

▪ Пример 5.5.5. — Распределение Коши.

Обозначение: $\boxed{\xi \sim C(x_0, \gamma)}$

Определение 5.5.5. Случайная величина ξ имеет ***распределение Коши с параметрами*** $\boldsymbol{x_0}$, $\boldsymbol{\gamma}$, если ее плотность распределения имеет вид

$$f_\xi(x) = \frac{1}{\pi}\cdot\frac{\gamma}{\gamma^2 + (x-x_0)^2}\,. \tag{5.25}$$

Рассмотрим следующий случайный эксперимент. Орудие закреплено на вращающейся в горизонтальной плоскости с постоянной угловой скоростью платформе. Центр вращения O находится на расстоянии γ от бесконечной стены. Ось Ox направлена вдоль стены, x_0 — проекция центра вращения на ось Ox. В случайный момент времени, когда линия прицела пересекает линию стены, происходит выстрел. Предполагая, что снаряд летит по линии прицела, обозначим через ξ координату точки попадания снаряда в стену. Тогда $\xi \sim C(x_0, \gamma)$.

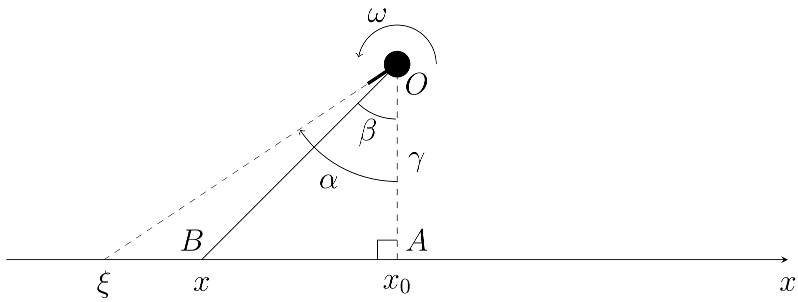


Рис. 5.5. Случайный эксперимент из примера 5.5.5

Действительно, пусть α — угол между траекторией снаряда (линией прицела) и перпендикуляром, опущенным из центра вращения на ось Ox (как обычно, положительным направлением отсчета угла считаем направление против часовой стрелки). Из условий случайного эксперимента следует, что $\alpha \sim U\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ (см. рис. 5.5). При этом

$$F_\xi(x) = P(\xi < x) = P(\alpha < \beta) = P\left(\alpha < \operatorname{arctg}\frac{x}{\gamma}\right) = \frac{\beta + \pi/2}{\pi}.$$

Находя угол β из треугольника OAB, получаем

$$F_\xi(x) = \frac{1}{\pi}\left(\operatorname{arctg}\frac{x - x_0}{\gamma} + \frac{\pi}{2}\right).$$

Дифференцируя $F_\xi(x)$, получаем (5.25). ■

5.6. Задачи для самостоятельного решения

1. Опишите $\sigma(\mathcal{K})$, если $\mathcal{K} = \big\{(-\infty; 0), (-\infty; 1)\big\}$.
2. Урна содержит 5 шаров с номерами от 1 до 5. Пусть ξ – наибольший номер, полученный при двух поштучных извлечениях с возвращением. Найти распределение вероятностей ξ. Найти вероятность того, что ξ примет значение, отличающееся от наиболее вероятного не больше, чем на 1.
3. Стрелок ведет огонь по удаляющейся мишени. Найти распределение вероятностей числа попаданий, если вероятность попадания при первом выстреле равна $0,6$, а при каждом последующем

выстреле она уменьшается на $0,2$. Найти вероятность того, что число попаданий будет не меньше двух.

4. На складе магазина имеется m экземпляров некоторого товара. Каждый посетитель магазина независимо от других покупает один экземпляр товара с вероятностью p. Найти распределение вероятностей номера посетителя, который купит последний экземпляр товара.
5. Точку наудачу бросают в квадрат $[0;1]\times[0;1]$ координатной плоскости. Пусть ξ — сумма, а η — разность ее абсциссы и ординаты. Найти функции распределения и плотности распределения случайных величин ξ и η.
6. Точку наудачу бросают в квадрат $[0;1]\times[0;1]$ координатной плоскости. Построить функции распределения и плотности распределения случайных величин φ и ρ — ее полярных координат. Найти $P\left(\dfrac{\pi}{6}<\varphi<\dfrac{\pi}{3}\right)$ и $P\left(\dfrac{1}{2}<\rho<\dfrac{2}{\sqrt{3}}\right)$, используя функции распределения.
7. В условиях предыдущей задачи построить функцию распределения вероятностей и плотность распределения вероятностей случайной величины θ, определенной как произведение абсциссы и ординаты точки. С помощью функции распределения найти $P\left(\dfrac{1}{2}<\theta<\dfrac{3}{4}\right)$.
8. Найти функцию распределения и плотность распределения вероятностей случайной величины $\eta=\xi^2$, если

   a) $\xi\sim\mathrm{Exp}(1)$;

   b) $\xi\sim U(-1,1)$.
9. Пусть $\xi\sim U(0,2\pi)$. Найти функцию распределения и плотность распределения вероятностей случайной величины $\eta=|\sin\xi|$.

Глава 6. Случайные векторы. Совместные распределения

При построении вероятностных моделей часто возникает необходимость рассматривать несколько случайных величин, определенных на одном вероятностном пространстве. Таким образом, возникает понятие случайного вектора.

Определение 6.0.1. Упорядоченная n-ка случайных величин, определенных на некотором вероятностном пространстве $(\Omega, \mathcal{F}, P)$, называется n-мерным ***случайным вектором***:

$$\vec{\xi} = (\xi_1, \xi_2, \ldots, \xi_n)^T = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}.$$

Случайный вектор является отображением из Ω в пространство $\mathbb{R}^n$. В дальнейшем мы будем всегда считать $\mathbb{R}^n$ евклидовым векторным пространством столбцов $\overline{x} = (x_1, x_2, \ldots, x_n)^T$ действительных чисел. Можно доказать, что $\vec{\xi} : \Omega \to \mathbb{R}^n$ является $\mathcal{F}|\mathcal{B}(\mathbb{R}^n)$-измеримым[1] отображением, т. е. для любого $B \in \mathcal{B}(\mathbb{R}^n)$

$$\vec{\xi}^{-1}(B) = \left\{\omega \in \Omega \,|\, \vec{\xi}(\omega) = (\xi_1(\omega), \ldots, \xi_n(\omega))^T \in B\right\} \in \mathcal{F},$$

[1] Через $\mathcal{B}(\mathbb{R}^n)$ обозначаем сигма-алгебру борелевских подмножеств пространства $\mathbb{R}^n$, которая определяется как минимальная сигма-алгебра подмножеств $\mathbb{R}^n$, содержащая все n-мерные параллелепипеды $\Pi = \bigotimes\limits_{k=1}^{n} [a_k; b_k)$, где $a_k, b_k \in \mathbb{R}$.

т. е. множество $\vec{\xi}^{-1}(B)$ является событием.

Определение 6.0.2. Вероятностная мера $P_{\vec{\xi}} = P_{\xi_1,\ldots,\xi_n} : \mathcal{B}(\mathbb{R}^n) \to [0;1]$, определенная равенством

$$\forall\ B \in \mathcal{B}(\mathbb{R}^n) \quad P_{\vec{\xi}}(B) = P_{\xi_1,\ldots,\xi_n}(B) := P\left(\vec{\xi} \in B\right), \tag{6.1}$$

называется ***распределением вероятностей случайного вектора*** $\vec{\xi} = (\xi_1, \xi_2, ..., \xi_n)^T$, или ***совместным распределением вероятностей*** случайных величин $\xi_1, \ldots, \xi_n$. Распределения вероятностей компонент случайного вектора называют ***маргинальными*** (по отношению к совместному распределению вероятностей).

Рассмотрим способы описания совместных распределений вероятностей случайных величин.

6.1. Совместные распределения вероятностей дискретных случайных величин

Рассмотрим дискретный случайный вектор $\vec{\xi} = (\xi_1, \xi_2, \ldots, \xi_n)^T$, где $\xi_1, \ldots, \xi_n$ — дискретные случайные величины, определенные на вероятностном пространстве $(\Omega, \mathcal{F}, P)$:

$$\xi_i(\omega) = \sum_j x_j^i \mathbf{1}_{A_j^i}(\omega), \quad A_j^i \in \mathcal{F}, \quad P(A_j^i) = P(\xi = x_j^i) = p_j^i. \tag{6.2}$$

Напомним, что мы по умолчанию полагаем $x_{j_1}^i \neq x_{j_2}^i$ при $j_1 \neq j_2$ для любых $i = 1, \ldots, n$.

Предложение 6.1.1. Совместное распределение вероятностей дискретных случайных величин $\xi_1, \ldots, \xi_n$ вида (6.2) полностью определяется набором вероятностей

$$p_{j_1,j_2,\ldots,j_n} = P\big(\xi_1 = x_{j_1}^1, \xi_2 = x_{j_2}^2, \ldots, \xi_n = x_{j_n}^n\big).$$

Действительно, для любого $B \in \mathcal{B}(\mathbb{R}^n)$ имеем:

$$P\left(\vec{\xi} \in B\right) = P\left(\sum_{(x^1_{j_1}, x^2_{j_2}, \ldots, x^n_{j_n})^T \in B} \{\xi_1 = x^1_{j_1}, \xi_2 = x^2_{j_2}, \ldots, \xi_n = x^n_{j_n}\}\right) = \\ = \sum_{(x^1_{j_1}, x^2_{j_2}, \ldots, x^n_{j_n})^T \in B} p_{j_1, j_2, \ldots, j_n},$$

так как события вида $\{\xi_1 = x^1_{j_1}, \xi_2 = x^2_{j_2}, \ldots, \xi_n = x^n_{j_n}\}$, соответствующие различным n-кам $(x^1_{j_1}, x^2_{j_2}, \ldots, x^n_{j_n})^T$, попарно несовместны.

Для случайного вектора $\vec{\xi} = (\xi_1, \ldots, \xi_n)^T$ с дискретными компонентами вида (6.2) маргинальные распределения вероятностей строятся при помощи формулы

$$p^i_j = \sum_{j_1, \cdots < j_i > \ldots, j_n} p_{j_1, \ldots, \underset{i}{j}, \ldots, j_n}. \tag{6.3}$$

Символ $j_1, \cdots < j_i > \ldots, j_n$ под знаком суммы означает, что суммирование ведется по всем возможным значениям всех индексов, кроме i-го. При этом индекс j_i полагается равным j.

R В двумерном случае при описании маргинальных распределений вероятностей компонент дискретного случайного вектора часто используют обозначения $p_{i\cdot}$ и $p_{\cdot j}$:

$$p_{i\cdot} := p^1_i = \sum_j p_{ij}, \quad p_{\cdot j} := p^2_j = \sum_i p_{ij}.$$

■ Пример 6.1.1. Игральную кость бросают до первого появления числа меньшего пяти. Пусть ξ — число бросков, η — число очков при последнем броске. Опишем распределение вероятностей случайного вектора $\vec{\xi} = (\xi, \eta)^T$. Пусть

$$B_k = \{\text{при } k\text{-м броске выпало 5 или 6}\},$$
$$A_{ij} = \{\text{при } i\text{-м броске выпало } j\}.$$

Рассмотрим событие $\{\xi = i, \eta = j\}$, где $i \in \mathbb{N}$, $j \in \{1; 2; 3; 4\}$. Оно происходит тогда, когда до i-го броска выпадали пятерка или шестерка, а при i-м броске выпало число j. Поэтому

$$p_{ij} = P\left(\xi = i, \eta = j\right) = P\left(B_1 \dots B_{i-1} A_{ij}\right) = \\ = P(B_1) \dots P(B_{i-1}) P(A_{ij}) = \left(\frac{1}{3}\right)^{i-1} \frac{1}{6}\,.$$

Здесь мы воспользовались независимостью событий $B_1, \dots, B_{i-1}$ и A_{ij}.

Найдем маргинальные распределения вероятностей. Для распределения вероятностей случайной величины ξ имеем:

$$p_{i\,\cdot} = \sum_j p_{ij} = \sum_{j=1}^{4} \left(\frac{1}{3}\right)^{i-1} \frac{1}{6} = \frac{2}{3}\left(\frac{1}{3}\right)^{i-1}\,.$$

Это согласуется с тем, что из условия задачи очевидно следует, что $\xi \sim \mathrm{Geom}\left(\frac{2}{3}\right)$. Для распределения вероятностей случайной величины η получаем:

$$p_{\cdot j} = \sum_i p_{ij} = \sum_{i=1}^{\infty} \left(\frac{1}{3}\right)^{i-1} \frac{1}{6} = \frac{1}{6} \cdot \frac{1}{1 - \frac{1}{3}} = \frac{1}{4}\,.$$

Это согласуется с тем, что из условия задачи можно понять, что $\eta \sim U\{1; 2; 3; 4\}$. ■

6.2. Функция совместного распределения

В общем случае для описания совместного распределения вероятностей случайных величин используют функцию совместного распределения.

Определение 6.2.1. Функция распределения вероятностей случайного вектора $\vec{\xi} = (\xi_1, \dots, \xi_n)^T$, или ***функция совместного распределения*** случайных величин $\xi_1, \dots, \xi_n$ определяется равенством

$$F_{\vec{\xi}}(\overline{x}) = F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) := P\left(\xi_1 < x_1, \dots, \xi_n < x_n\right)\,,$$

для любого $\overline{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n\,.$

Функции маргинальных распределений при этом находятся по формуле

$$F_{\xi_i}(x) = \lim_{\substack{x_1\to+\infty\\ \cdots\\ x_i=x\\ \cdots\\ x_n\to+\infty}} F_{\xi_1,\ldots,\xi_n}(x_1,\ldots,x_n)\,. \tag{6.4}$$

Действительно, рассмотрим последовательность событий

$$A_k = \{\xi_1 < k, \ldots, \xi_{i-1} < k, \xi_i < x, \xi_{i+1} < k, \ldots, \xi_n < k\}\ , \quad k \in \mathbb{N}\,.$$

Для любого $k \in \mathbb{N}$ имеют место вложения $A_k \subseteq A_{k+1}$, при этом $\sum\limits_{k\in\mathbb{N}} A_k = \{\xi_i < x\}$, поэтому (по свойству непрерывности вероятности (P-9)), получаем

$$F_{\xi_i}(x) = P(\xi_i < x) = P\left(\sum_{k\in\mathbb{N}} A_k\right) = \lim_{k\to\infty} P(A_k) =$$

$$= \lim_{k\to\infty} F_{\xi_1,,\xi_n}(k,\ldots,k,x,k,\ldots,k) = \lim_{\substack{x_1\to+\infty\\ \cdots\\ x_i=x\\ \cdots\\ x_n\to+\infty}} F_{\xi_1,\ldots,\xi_n}(x_1,\ldots,x_n)\,.$$

▪ Пример 6.2.1. Пусть ξ и η — моменты прихода участников эксперимента в задаче о встрече (см. пример 5.4.2). Рассмотрим случайный вектор $\vec{\xi} = (\xi,\eta)^T$. Его функция распределения (функция совместного распределения ξ и η) имеет вид:

$$F_{\vec{\xi}}(\overline{x}) = F_{\xi\eta}(x,y) = \begin{cases} 0\,, & x<0\ \vee\ y<0\,,\\ xy\,, & 0\le x<1\ \wedge\ 0\le y<1\,,\\ x\,, & 0\le x<1\ \wedge\ y\ge 1\,,\\ y\,, & x\ge 1\ \wedge\ 0\le y<1\,,\\ 1\,, & x\ge 1\wedge y\ge 1\,. \end{cases} \tag{6.5}$$

Проверьте это самостоятельно. ▪

6.3. Совместное распределение абсолютно непрерывных случайных величин

Определение 6.3.1. Случайный вектор $\vec{\xi} = (\xi_1, \ldots, \xi_n)^T$ называется ***абсолютно непрерывным***, если его распределение вероятностей $P_{\vec{\xi}}$ имеет ***плотность***, т. е. существует функция $f_{\vec{\xi}} = f_{\xi_1,\ldots,\xi_n} : \mathbb{R}^n \to \mathbb{R}$ такая, что для любого $\overline{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$

$$F_{\vec{\xi}}(\overline{x}) = F_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n) = \\ = \int\limits_{-\infty}^{x_1} \ldots \int\limits_{-\infty}^{x_n} f_{\vec{\xi}}(\overline{x})\, d\overline{x} = \int\limits_{-\infty}^{x_1} \ldots \int\limits_{-\infty}^{x_n} f_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n)\, dx_1 \ldots dx_n\,.$$

Из этого определения и формулы (6.4) следует формула для плотности маргинального распределения вероятностей:

$$f_{\xi_i}(x) = \underset{\mathbb{R}^{n-1}}{\int \cdots \int} f_{\xi_1,\ldots,\xi_n}(x_1, \ldots, \underset{i}{x}, \ldots, x_n)\, dx_1 \ldots <dx_i> \ldots dx_n\,. \qquad (6.6)$$

Эта формула является континуальным аналогом дискретной формулы (6.3), где роль индексов $j_1, \ldots, j_n$, пробегающих дискретные множества значений, играют переменные $x_1, \ldots, x_n$, пробегающие $\mathbb{R}$, а вместо суммирования используется интегрирование.

Плотность распределения вероятностей случайного вектора обладает свойствами, аналогичными свойствам плотности распределения вероятностей одномерной случайной величины. Приведем без доказательства многомерный аналог предложения 5.5.1.

Предложение 6.3.1. Если $\vec{\xi}$ — абсолютно непрерывный n-мерный случайный вектор, то его функция распределения $F_{\vec{\xi}}$ непрерывна во всех точках $\mathbb{R}^n$, а его плотность распределения $f_{\vec{\xi}}$ обладает свойствами

(ff-I) $\forall\ \overline{x} \in \mathbb{R}^n \quad f_{\vec{\xi}}(\overline{x}) \geq 0\,;$

(ff-II) $\underset{\mathbb{R}^n}{\int \cdots \int} f_{\vec{\xi}}(\overline{x})\, d\overline{x} = 1\,;$

(ff-III) В точках непрерывности функции $f_{\vec{\xi}}$ функция распределения $F_{\vec{\xi}}$ n раз дифференцируема и справедливо равенство:

$$\frac{\partial^n}{\partial x_1 \ldots \partial x_n} F_{\vec{\xi}}(x_1, \ldots, x_n) = f_{\vec{\xi}}(x_1, \ldots, x_n)\,. \tag{6.7}$$

(ff-IV) $\forall\ B \in \mathcal{B}(\mathbb{R}^n):\quad P(\vec{\xi} \in B) = \underset{B}{\int \cdots \int} f_{\vec{\xi}}(\overline{x})\, d\overline{x}\,.$

▪ Пример 6.3.1. Плотность распределения $f_{\xi\eta}$ случайного вектора $\vec{\xi} = (\xi, \eta)^T$ (плотность совместного распределения ξ и η) из примера 6.2.1 находится в точках непрерывности $f_{\xi\eta}$ по функции $F_{\xi\eta}$ совместного распределения ξ и η, заданной равенством (6.5), с помощью формулы (6.7). Дифференцируя, получаем:

$$f_{\xi\eta}(x, y) = \begin{cases} 1\,, & (x, y)^T \in (0; 1) \times (0; 1)\,, \\ 0\,, & (x, y)^T \notin (0; 1) \times (0; 1)\,. \end{cases}$$

В точках разрыва, образующих границу квадрата $[0; 1] \times [0; 1]$, плотность $f_{\xi\eta}$ можно доопределить произвольным образом — это не влияет на распределение вероятностей. Поэтому окончательно можно положить

$$f_{\xi\eta}(x, y) = \mathbf{1}_{[0;1]\times[0;1]}(x, y)\,.$$ ▪

Определение 6.3.2. Говорят, что случайный вектор $\vec{\xi} = (\xi_1 \ldots, \xi_n)^T$ ***равномерно распределен в области $D \subset \mathbb{R}^n$***, если

$$f_{\vec{\xi}}(\overline{x}) = \frac{1}{\mu(D)} \cdot \mathbf{1}_D(\overline{x})\,,$$

где μ — мера Лебега в $\mathbb{R}^n$. При этом пишут: $\vec{\xi} \sim U(D)$.

Таким образом, пример 6.3.1 показывает, что совместное распределение моментов прихода участников эксперимента, описанного в задаче о встрече, является равномерным в квадрате $[0; 1] \times [0; 1]$.

Приведем многомерный аналог предложения 5.5.2.

Предложение 6.3.2. Пусть $\vec{\eta}$ — n-мерный случайный вектор, определенный на вероятностном пространстве $(\Omega, \mathcal{F}, P)$, $f_{\vec{\eta}}(\overline{x})$ — плотность распределения вероятностей $\vec{\eta}$, $\overline{g} : \mathbb{R}^n \to \mathbb{R}^n$ — непрерывно-дифференцируемое биективное отображение. Тогда плотность распределения вероятностей случайного вектора $\vec{\xi} = \overline{g}(\vec{\eta})$ имеет вид

$$f_{\vec{\xi}}(\overline{x}) = \frac{1}{|J(\overline{g})|} f_{\vec{\eta}}\big(\overline{g}^{-1}(\overline{x})\big)\,, \tag{6.8}$$

где $J(\overline{g})$ — матрица Якоби отображения $\overline{g}$, а $|J(\overline{g})|$ — абсолютная величина его якобиана.

Доказательство. Для любого $B \in \mathcal{B}(\mathbb{R}^n)$ справедливо равенство

$$P\big(\vec{\xi} \in B\big) = \int\limits_B f_{\vec{\xi}}(\overline{x})\, d\overline{x}\,.$$

С другой стороны, делая замену переменных в кратном интеграле, получаем:

$$P\big(\vec{\xi} \in B\big) = P\big(\vec{\eta} \in \overline{g}^{-1}(B)\big) = \int\limits_{\overline{g}^{-1}(B)} f_{\vec{\eta}}(\overline{y})\, d\overline{y} =$$

$$= \left[\begin{array}{c} \overline{y} = \overline{g}^{-1}(\overline{x}) \\ d\overline{y} = \left|J(\overline{g}^{-1})\right|\, d\overline{x} = \dfrac{d\overline{x}}{|J(\overline{g})|} \end{array} \right] = \int\limits_B f_{\vec{\eta}}\big(\overline{g}^{-1}(\overline{x})\big) \frac{d\overline{x}}{|J(\overline{g})|}\,.$$

Таким образом, получаем равенство

$$\int\limits_B f_{\vec{\xi}}(\overline{x})\, d\overline{x} = \int\limits_B f_{\vec{\eta}}\big(\overline{g}^{-1}(\overline{x})\big) \frac{d\overline{x}}{|J(\overline{g})|}\,.$$

В силу того, что оно верно для любого $B \in \mathcal{B}(\mathbb{R}^n)$, из него следует равенство подинтегральных функций, т. е. равенство (6.8). ∎

6.4. Независимость случайных величин

Пусть случайная величина ξ определена на вероятностном пространстве $(\Omega, \mathcal{F}, P)$.

Определение 6.4.1. Сигма-алгебра подмножеств пространства элементарных событий Ω, определенная равенством

$$\sigma(\xi) := \left\{ \xi^{-1}(B) \,|\, B \in \mathcal{B}(\mathbb{R}) \right\} , \tag{6.9}$$

называется ***сигма-алгеброй, порожденной случайной величиной $\boldsymbol{\xi}$***.

Сигма-алгебра, порожденная случайной величиной ξ, представляет собой совокупность всех событий вида $\{\xi \in B\}$, где $B \in \mathcal{B}(\mathbb{R})$. Фактически это совокупность всех событий, которые могут произойти со случайной величиной ξ в результате рассматриваемого случайного эксперимента.

▪ Пример 6.4.1. Пусть ξ и η — моменты прихода участников эксперимента в задаче о встрече (см. пример 5.4.2), $\zeta = \eta - \xi$. Порожденные этими случайными величинами сигма-алгебры допускают явное описание:

$$\sigma(\xi) = \left\{ B \times [0; 1] \,|\, B \in \mathcal{B}(\mathbb{R}) \right\} ,$$
$$\sigma(\eta) = \left\{ [0; 1] \times B \,|\, B \in \mathcal{B}(\mathbb{R}) \right\} ,$$
$$\sigma(\zeta) = \left\{ \{\omega = (u; v) \in \Omega \,|\, u + v \in B\} \,|\, B \in \mathcal{B}(\mathbb{R}) \right\} .$$

На рис. 6.1 изображены элементы этих сигма-алгебр, соответствующие множеству $B = [a; b]$. ▪

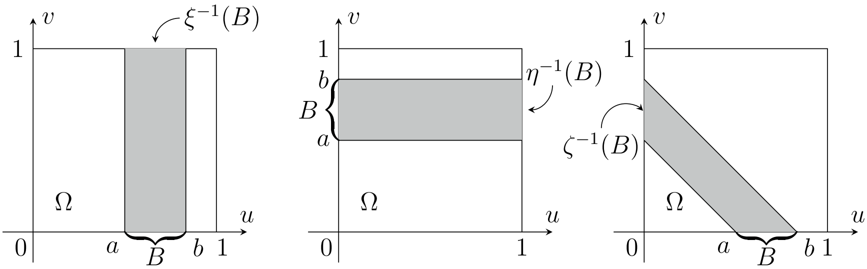


Рис. 6.1. Примеры элементов сигма-алгебр $\sigma(\xi)$, $\sigma(\eta)$ и $\sigma(\zeta)$

Определение 6.4.2. Пусть $(\Omega, \mathcal{F}, P)$ — вероятностное пространство. ***Сигма-алгебры*** $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_n$, где $\mathcal{F}_i \subseteq \mathcal{F}$, $i = 1, 2, \ldots, n$, называются ***независимыми***, если

$$\forall\ A_1 \in \mathcal{F}_1, \ldots, A_n \in \mathcal{F}_n \quad P\left(\prod_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i)\,. \tag{6.10}$$

R Независимость сигма-алгебр $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_n$ означает независимость в совокупности любого набора событий $A_1, A_2, \ldots, A_n$, взятых по одному из каждой сигма-алгебры: $A_i \in \mathcal{F}_i$, $i = 1, 2, \ldots, n$. Действительно (см. определение 2.3.3), для любого набора индексов i_k таких, что $1 \le i_1 < i_2 < \ldots, i_p \le n$, выполняется равенство

$$P(A_{i_1} A_{i_2} \ldots A_{i_p}) = P(A_1 A_2 \ldots A_n)\,,$$

где $A_i = A_{i_k}$, если $i = i_k\,, k = 1, \ldots, p$ и $A_i = \Omega$, если $i \notin \{i_1, \ldots, i_p\}$. Так как каждая из сигма-алгебр $\mathcal{F}_i$ содержит Ω, далее можем воспользоваться их независимостью и равенством (6.10), в результате получим

$$P(A_{i_1} A_{i_2} \ldots A_{i_p}) = P(A_1)P(A_2)\ldots P(A_n) = P(A_{i_1})P(A_{i_2})\ldots P(A_{i_p})\,,$$

так как $P(A_i) = 1$, если $A_i = \Omega$.

Определение 6.4.3. ***Случайные величины*** $\xi_1, \xi_2, \ldots, \xi_n$ называются ***независимыми***, если независимы порожденные ими сигма-алгебры $\sigma(\xi_1), \sigma(\xi_2), \ldots, \sigma(\xi_n)$.

■ Пример 6.4.2. Рассмотрим случайные величины ξ и η из примера 6.4.1. Они независимы, так как для любых $B_1, B_2 \in \mathcal{B}(\mathbb{R})$ события $\{\xi \in B_1\}$ и $\{\eta \in B_2\}$ независимы. Действительно,

$$P\left(\{\xi \in B_1\} \cdot \{\eta \in B_2\}\right) = \mu(B_1 \times B_2)\,,$$

где μ — мера Лебега на плоскости (площадь). Но $\mu(B_1 \times B_2) = \mu(B_1)\cdot$ $\cdot\mu(B_2)$, где $\mu(B_i)$ — мера Лебега множества B_i в $\mathbb{R}$ (длина), так что

$$P\left(\{\xi \in B_1\} \cdot \{\eta \in B_2\}\right) = \mu(B_1) \cdot \mu(B_2) =$$
$$= \mu\left(B_1 \times [0;1]\right) \cdot \mu\left([0;1] \times B_2\right) = (\xi \in B_1) \cdot P(\eta \in B_2)\,.$$

Независимость случайных величин «на интуитивном уровне» означает, что если у нас есть информация о том, какое значение приняла одна из них (в частности, какие из событий, принадлежащих порожденной ею сигма-алгебре, произошли, а какие — нет), это не меняет шансы событий, которые могут произойти с другой случайной величиной.

Для $B = [a;b] \subset [0;1]$ (см. рис. 6.2) события $\{\xi \in B\}$ и $\{\zeta \in B\}$ не являются независимыми. Действительно,

$$P(\zeta \in B) = (b^2 - a^2)/2\,, \quad P(\xi \in B) = b - a\,,$$

при этом

$$P\big(\{\xi \in B\} \cdot \{\zeta \in B\}\big) = \frac{(b-a)^2}{2}\,,$$

так что $P\big(\{\xi \in B\} \cdot \{\zeta \in B\}\big) \neq P(\xi \in B) \cdot P(\zeta \in B)$. Таким образом, случайные величины ξ и ζ зависимы. Аналогично, зависимыми являются η и ζ.

Зависимость ξ и ζ в этом примере интуитивно понятна, так как информация о том, какое значение приняла ζ, меняет шансы произойти для событий вида $\{\xi \in B\}$, где $B \in \mathcal{B}(\mathbb{R})$. Например, если ζ приняла значение большее $\frac{3}{2}$, то ξ уже не может принять значения меньшего $\frac{1}{2}$. ∎

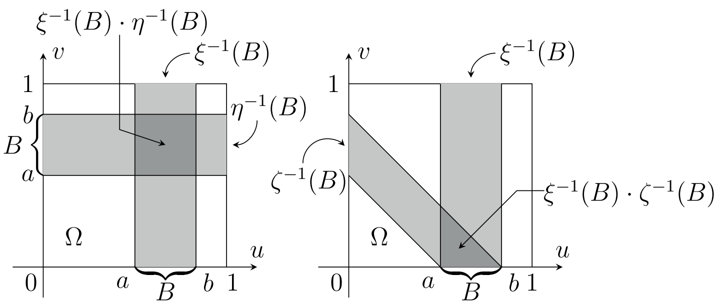


Рис. 6.2. Независимость сигма-алгебр $\sigma(\xi)$ и $\sigma(\eta)$, зависимость сигма-алгебр $\sigma(\xi)$ и $\sigma(\zeta)$

6.5. Критерии независимости случайных величин

На практике при работе со случайными величинами далеко не всегда можно дать явное описание сигма-алгебры, порожденной той или иной случайной величиной. Рассмотрим критерии независимости случайных величин в терминах соотношений, связывающих их совместные распределения вероятностей с маргинальными распределениями вероятностей.

Теорема 6.5.1. — Критерий независимости дискретных случайных величин. Пусть совместное распределение дискретных случайных величин $\xi_1, \ldots, \xi_n$ задано вероятностями

$$p_{j_1,j_2,\ldots,j_n} = P\big(\xi_1 = x^1_{j_1}, \xi_2 = x^2_{j_2}, \ldots, \xi_n = x^n_{j_n}\big)\,,$$

а маргинальные распределения вероятностей заданы вероятностями $P(\xi = x^i_j) = p^i_j$. Случайные величины $\xi_1, \ldots, \xi_n$ независимы тогда и только тогда, когда

$$p_{j_1,j_2,\ldots,j_n} = \prod_{i=1}^{n} p^i_{j_i} \tag{6.11}$$

для любых $j_1, j_2, \ldots, j_n$.

Доказательство. Необходимость условия (6.11) очевидна. Для доказательства достаточности возьмем $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$ и рассмотрим события $\{\xi_i \in B_i\} \in \sigma(\xi_i)\,, i = 1, \ldots, n$. Для них имеем:

$$\begin{aligned}
P\left(\prod_{i=1}^{n}\{\xi_i \in B_i\}\right) &= P\left(\sum_{x^1_{j_1}\in B_1,\ldots,x^n_{j_n}\in B_n} \{\xi_1 = x^1_{j_1}\}\cdot\ldots\cdot\{\xi_n = x^n_{j_n}\}\right) = \\
&= \sum_{x^1_{j_1}\in B_1,\ldots,x^n_{j_n}\in B_n} P\left(\{\xi_1 = x^1_{j_1}\}\cdot\ldots\cdot\{\xi_n = x^n_{j_n}\}\right) = \\
&= \sum_{x^1_{j_1}\in B_1,\ldots,x^n_{j_n}\in B_n} p_{j_1,j_2,\ldots,j_n} = \sum_{x^1_{j_1}\in B_1,\ldots,x^n_{j_n}\in B_n} \prod_{i=1}^{n} p^i_{j_i} = \\
&= \prod_{i=1}^{n} \sum_{x^i_{j_i}\in B_i} p^i_{j_i} = \prod_{i=1}^{n} P\left(\xi_i \in B_i\right)\,,
\end{aligned}$$

т. е. для любых $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$ события $\{\xi_i \in B_i\}\,, i = 1, \ldots, n$ независимы. Это означает независимость сигма-алгебр $\sigma(\xi_i)$. ■

Теорема 6.5.2. — Критерий независимости случайных величин в терминах функций распределения.

Пусть совместное распределение вероятностей случайных величин $\xi_1, ..., \xi_n$ задано функцией совместного распределения $F_{\vec{\xi}} = F_{\xi_1,...,\xi_n}$, а F_{ξ_i} , $i = 1, ..., n$ — функции маргинальных распределений вероятностей. Тогда для независимости случайных величин $\xi_1, ..., \xi_n$ необходимо и достаточно, чтобы выполнялось условие

$$F_{\xi_1,...,\xi_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} F_{\xi_i}(x_i)\,. \tag{6.12}$$

Необходимость условия (6.12) очевидна. Достаточность — нетривиальный факт, который мы примем без доказательства (его можно найти, например в книге [8]).

Теорема 6.5.3. — Критерий независимости в терминах плотностей распределения вероятностей.

Пусть совместное распределение вероятностей случайных величин $\xi_1, ..., \xi_n$ задано плотностью $f_{\vec{\xi}} = f_{\xi_1,....,\xi_n}$, а f_{ξ_i} , $i = 1, ..., n$ — плотности маргинальных распределений вероятностей. Тогда для независимости случайных величин $\xi_1, ..., \xi_n$ необходимо и достаточно, чтобы выполнялось условие

$$f_{\xi_1,...,\xi_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_{\xi_i}(x_i)\,. \tag{6.13}$$

Теорема является следствием теоремы 6.5.2 в силу равенства (6.7).

6.6. Задачи для самостоятельного решения

1. Игральную кость, у которой на двух гранях находятся единицы, на двух — двойки, а на оставшихся — тройки, бросают до первого появления тройки. Пусть ξ — число бросков, η — индикатор четности суммы очков, полученной в результате. Найти совместное распределение вероятностей случайных величин ξ и η и их маргинальные распределения. Являются ли ξ и η независимыми?

2. В N ячеек независимо бросают частицы. Каждая частица с одинаковой вероятностью может попасть в любую из них (вместимость ячеек не ограничена). Пусть $\nu_1 < \nu_2 < \ldots < \nu_N$ — номера бросков, при которых частицы попадают в пустые ячейки, θ_k — номер ячейки, в которую попадает частица при ν_k-м броске. Найти совместное распределение вероятностей случайных величин $\theta_1, \theta_2, \theta_3$ и их маргинальные распределения. Являются ли эти случайные величины независимыми?
3. В условиях предыдущей задачи пусть $\tau_k = \nu_k - \nu_{k-1}\,, k \geq 2$. Найти совместное распределение случайных величин τ_2 и τ_3 и их маргинальные распределения. Являются ли эти случайные величины независимыми?
4. Случайный вектор $\vec{\xi} = (\xi, \eta)^T \sim U(S)$, где
$$S = \left\{(x; y) \in \mathbb{R}^2 \,|\, x^2 + y^2 \leq 1\right\}.$$
Найти плотность совместного распределения случайных величин φ и ρ — полярных координат $\vec{\xi}$. По совместному распределению восстановить маргинальные распределения вероятностей. Являются ли φ и ρ независимыми?
5. Пусть $\vec{\xi} = (\xi, \eta)^T \sim N(\overline{0}, E)$ — двумерный сферически-симметрический нормальный вектор (см. Приложение C), φ и ρ — полярные координаты $\vec{\xi}$. Построить плотность совместного распределения вероятностей φ и $\theta = \rho^2$, восстановить по нему маргинальные распределения. Являются ли φ и θ независимыми?
6. Случайные величины $\xi \sim \mathrm{Exp}(1)$ и $\eta \sim \mathrm{Exp}(1)$ независимы. Найти плотность совместного распределения вероятностей случайных величин $\zeta = \xi + \eta$ и $\nu = \xi - \eta$, их маргинальные распределения. Являются ли ζ и ν независимыми?

Глава 7. Моменты случайных величин

Моменты являются важными числовыми характеристиками распределений вероятностей случайных величин. Базовой характеристикой является момент первого порядка, который называется математическим ожиданием случайной величины.

7.1. Математическое ожидание

Математическое ожидание случайной величины ξ обозначается $E\xi$ по первой букве слова *expectation*. Определяется оно в два этапа: на первом — для дискретных случайных величин, на втором — с помощью предельного перехода определение распространяется на максимально широкий класс случайных величин.

Математическое ожидание дискретной случайной величины

Определение 7.1.1. Пусть дискретная случайная величина ξ определена на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ равенством

$$\xi(\omega) = \sum_k x_k \mathbf{1}_{A_k}(\omega)\,, \quad x_k \in \mathbb{R}\,,\ A_k \in \mathcal{F}\,, \quad P(A_k) = p_k\,. \qquad (7.1)$$

Математическое ожидание случайной величины ξ определяется равенством

$$E\xi := \sum_k x_k p_k \,, \tag{7.2}$$

при условии, что ряд в правой части равенства (7.2) абсолютно сходится.

R Понятие математического ожидания в теории вероятностей является аналогом понятия «центр тяжести» для системы масс p_k, сосредоточенных в точках x_k числовой оси. Математическое ожидание представляет собой точку равновесия распределения вероятностей случайной величины ξ. При этом на само распределение вероятностей можно смотреть как на систему вероятностных масс p_k сосредоточенных в точках x_k. Математическое ожидание также называют средним значением случайной величины (*mean value*).

R Представление случайной величины ξ в виде разложения (7.1), как уже отмечалось, не является однозначным (см. замечание на стр. 66), однако нетрудно проверить, что определение математического ожидания дискретной случайной величины ξ формулой (7.2) не зависит от выбора представления ξ.

R Из определения следует, что математическое ожидание дискретной случайной величины ξ существует тогда и только тогда, когда выполнено условие $E|\xi| = \sum_k |x_k| p_k < \infty$.

Рассмотрим примеры вычисления математических ожиданий для некоторых из дискретных распределений вероятностей, которые были введены в главе 5.

■ Пример 7.1.1. Пусть $\xi \sim U\{x_1, x_2, \ldots, x_n\}$, тогда очевидно, что

$$E\xi = \frac{x_1 + x_2 + \cdots + x_n}{n} \,,$$

т. е. математическое ожидание (среднее значение) случайной величины, равномерно распределенной по значениям $x_1, x_2, \ldots, x_n$, совпадает со средним арифметическим этих значений. ■

■ Пример 7.1.2. Пусть $\xi \sim B(n,p)$, тогда $E\xi = np$.

Действительно, обозначая $q = 1 - p$, получаем

$$E\xi = \sum_{k=0}^{n} kC_n^k p^k q^{n-k} = p\sum_{k=1}^{n} kC_n^k p^{k-1} q^{n-k} = p \cdot \frac{d}{dp}\left(\sum_{k=0}^{n} C_n^k p^k q^{n-k}\right) =$$
$$= p \cdot \frac{d}{dp}\,(p+q)^n = np(p+q)^{n-1} = np\,.$$

Этот результат согласуется с интуитивным представлением о среднем числе успехов в серии испытаний Бернулли, если интерпретировать вероятность успеха p как среднюю долю испытаний, заканчивающихся успехом. ■

■ Пример 7.1.3. Пусть $\xi \sim P(\lambda)$, тогда $E\xi = \lambda$.

Действительно, имеем:

$$E\xi = \sum_{k=0}^{\infty} k\frac{\lambda^k}{k!}e^{-\lambda} = \lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = \lambda e^{-\lambda}e^{\lambda} = \lambda\,.$$

Этот результат согласуется с интерпретацией параметра λ в математической модели радиоактивного распада, рассмотренной в примере 5.3.4, где было показано, что количество частиц, испускаемых радиоактивным образцом за единицу времени, можно считать распределенным по закону Пуассона с параметром λ, где λ — среднее число частиц, испускаемых за единицу времени. ■

■ Пример 7.1.4. Пусть $\xi \sim \mathrm{Geom}(p)$, тогда $E\xi = \dfrac{1}{p}$.

Действительно, справедливы равенства:

$$E\xi = \sum_{k=1}^{\infty} kq^{k-1}p = p \cdot \frac{d}{dq}\sum_{k=0}^{\infty} q^k = p \cdot \frac{d}{dq}\frac{1}{1-q} = \frac{p}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p}\,. \quad ■$$

Ⓡ Не у всякой дискретной случайной величины есть математическое ожидание. Например, пусть случайная величина ξ принимает значения $x_k = 2^k$ с вероятностями $p_k = \dfrac{1}{2^k}$, где $k = 1, 2, \ldots$. Тогда $E\xi = \infty$, так как ряд $\sum\limits_k x_k p_k$ в данном случае расходится.

Предложение 7.1.1. Математическое ожидание дискретных случайных величин обладает следующими свойствами:

(E-I) Если $\xi = c$, где $c = Const$, то $E\xi = c$;

(E-II) $E(\alpha\xi + \beta\eta) = \alpha E\xi + \beta E\eta$ для любых случайных величин ξ и η и констант α и β;

(E-III) Если случайные величины ξ и η независимы, то $E(\xi\eta) = E\xi \cdot E\eta$;

(E-IV) Если $\xi \le \eta$, то $E\xi \le E\eta$, в частности, если $\xi \le c$, где $c \in \mathbb{R}$ — константа, то $E\xi \le c$;

(E-V) $|E\xi| \le E|\xi|$.

Доказательство. Свойство (E-I) очевидно. Докажем (E-II). Представим ξ и η в виде

$$\xi(\omega) = \sum_k x_k \mathbf{1}_{A_k}(\omega), \quad \eta(\omega) = \sum_k y_k \mathbf{1}_{A_k}(\omega), \tag{7.3}$$

где $x_k, y_k \in \mathbb{R}$, $A_k \in \mathcal{F}$, $P(A_k) = p_k$. Такое представление (с общим набором событий A_k) возможно благодаря неоднозначности представления дискретной случайной величины (см. замечание на стр. 66 после определения 5.3.1). В результате получаем:

$$E(\alpha\xi + \beta\eta) = \sum_k (\alpha x_k + \beta y_k) p_k = \\ = \alpha \sum_k x_k p_k + \beta \sum_k y_k p_k = \alpha E\xi + \beta E\eta.$$

Докажем (E-III). Пусть $p_{ij} = P\big(\xi = x_i, \eta = y_j\big)$. Случайные величины ξ и η независимы тогда и только тогда, когда $p_{ij} = p_i^1 p_j^2$, где $p_i^1 = P(\xi = x_i)$, $p_j^2 = P(\eta = y_j)$ — вероятности маргинальных распределений. В результате

$$E(\xi\eta) = \sum_{i,j} x_i y_j p_{ij} = \sum_{i,j} x_i y_j p_i^1 p_j^2 = \sum_i x_i p_i^1 \cdot \sum_j y_j p_j^2 = E\xi \cdot E\eta.$$

Для доказательства (E-IV) предположим, без ограничения общности, что ξ и η имеют вид (7.3). Тогда $x_k \le y_k$ для любого k и, следовательно,

$$E\xi = \sum_k x_k p_k \le \sum_k y_k p_k = E\eta\,.$$

Свойство (E-V) очевидно следует из известного свойства абсолютных величин сумм. ■

R Свойство (E-I) остается верным, если $\xi = c$ с вероятностью 1.

▪ Пример 7.1.5. Пусть $\xi \sim HG(M, N, n)$, тогда $E\xi = \dfrac{nM}{N}$.

Заметим, что вычислить математическое ожидание случайной величины, имеющей гипергеометрическое распределение, пользуясь определением, очень сложно. Однако если разложить ее в сумму более простых случайных величин и воспользоваться линейностью математического ожидания, вычисление упрощается. Обычно в роли вспомогательных простых случайных величин при таком подходе используют индикаторы подходящих событий, поэтому этот метод вычисления математического ожидания часто называют методом индикаторов. Вспомним, что гипергеометрическое распределение вероятностей с параметрами N, M, n имеет случайная величина ξ, определенная как число черных шаров при извлечении наудачу n шаров из совокупности в N шаров, из которых M — черные. Представим извлечение шаров как последовательное. Пусть $A_k = \{k -$ й шар черный$\}$. Положим $\xi_k = \mathbf{1}_{A_k}$ — случайная величина, принимающая значение 1, если k-й шар черный и 0, если нет (индикатор черного при извлечении k-го шара). Тогда $\xi = \sum\limits_{k=1}^{n} \xi_k$. Поскольку $E\xi_k = 1 \cdot P(A_k) + 0 \cdot P(\overline{A}_k) = \dfrac{M}{N}$ (см. задачу 2 из раздела 1.6), то по свойству линейности математического ожидания (E-II) получаем $E\xi = \sum\limits_{k=1}^{n} E\xi_k = \dfrac{nM}{N}$. ▪

Приведем здесь еще несколько формул, использующихся для вычисления математических ожиданий функций от дискретных случайных величин.

Предложение 7.1.2. Пусть ξ — дискретная случайная величина, определенная формулой (7.1), а $g : \mathbb{R} \to \mathbb{R}$ — некоторая функция. Тогда

$$Eg(\xi) = \sum_k g(x_k)p_k \, . \tag{7.4}$$

Доказательство. Действительно, $g(\xi)$ — дискретная случайная величина, принимающая значеняия $g(x_k)$ с вероятностями p_k, поэтому формула (7.4) сразу следует из определения 7.1.1. ■

Аналогично доказывается и следующее предложение.

Предложение 7.1.3. Пусть $\vec{\xi} = (\xi_1, \xi_2, \dots, \xi_n)^T$ — случайный вектор с дискретными компонентами, распределение вероятностей которого задано набором вероятностей

$$p_{j_1,j_2,\dots,j_n} = P(\xi_1 = x^1_{j_1}, \xi_2 = x^2_{j_2}, \dots, \xi_n = x^n_{j_n}) \, ,$$

а $g : \mathbb{R}^n \to \mathbb{R}$ — некоторая функция. Тогда

$$Eg(\xi_1, \xi_2, \dots, \xi_n) = \sum_{j_1,j_2,\dots,j_n} g(x^1_{j_1}, x^2_{j_2}, \dots, x^n_{j_n}) p_{j_1,j_2,\dots,j_n} \, . \tag{7.5}$$

Общее определение математического ожидания

Определение 7.1.2. Пусть ξ — случайная величина, определенная на вероятностном пространстве $(\Omega, \mathcal{F}, P)$, $\{\xi_n\}$ — последовательность дискретных случайных величин, определенных на этом же вероятностном пространстве, равномерно сходящаяся к ξ на Ω[a]. Если $E|\xi_n| < \infty$ для любого $n \in \mathbb{N}$, то

$$E\xi := \lim_{n\to\infty} E\xi_n \, . \tag{7.6}$$

[a]Напомним, что равномерная на Ω сходимость последовательности $\{\xi_n\}$ к случайной величине ξ эквивалентна тому, что $\lim\limits_{n\to\infty} \left[\sup\limits_{\omega\in\Omega} |\xi_n(\omega) - \xi(\omega)| \right] = 0$.

Предложения 7.1.4, 7.1.5 и 7.1.6, рассматриваемые далее, доказывают корректность определения 7.1.2.

Предложение 7.1.4. Для любой случайной величины ξ, определенной на вероятностном пространстве $(\Omega, \mathcal{F}, P)$, существует последовательность $\{\xi_n\}$ дискретных случайных величин, определенных на этом же вероятностном пространстве, равномерно сходящаяся к ξ на Ω.

Доказательство. Определим

$$\xi_n(\omega) := \sum_{k\in\mathbb{Z}} \frac{k}{n} \cdot \mathbf{1}_{\left\{\frac{k}{n}\le\xi<\frac{k+1}{n}\right\}}(\omega)\,, \quad \omega\in\Omega\,. \tag{7.7}$$

В силу того, что

$$\left\{\frac{k}{n} \le \xi < \frac{k+1}{n}\right\} = \left\{\omega\in\Omega \,\middle|\, \frac{k}{n} \le \xi(\omega) < \frac{k+1}{n}\right\} = \\ = \xi^{-1}\left(\left[\frac{k}{n};\frac{k+1}{n}\right)\right) \in \mathcal{F}\,,$$

равенство (7.7) действительно определяет дискретную случайную величину на вероятностном пространстве $(\Omega, \mathcal{F}, P)$. При этом справедлива оценка $\sup\limits_{\omega\in\Omega}|\xi(\omega)-\xi_n(\omega)| \le \dfrac{1}{n}$, так как по определению $\xi_n(\omega) = \dfrac{k}{n}$ тогда и только тогда, когда $\xi(\omega) \in \left[\dfrac{k}{n};\dfrac{k+1}{n}\right)$. Из этой оценки следует равномерная на Ω сходимость последовательности $\{\xi_n\}$ к ξ. ■

Предложение 7.1.5. Для любой последовательности дискретных случайных величин $\{\xi_n\}$, имеющих конечные математические ожидания, равномерно сходящейся к случайной величине ξ, предел (7.6) существует.

Доказательство. Пусть $n, m \in \mathbb{N}$, $n < m$, тогда справедлива следующая оценка:

$$|E\xi_n - E\xi_m| \le E|\xi_n - \xi_m| \le \sup_{\omega\in\Omega}|\xi_n(\omega)-\xi_m(\omega)| \le \\ \le \sup_{\omega\in\Omega}|\xi_n(\omega)-\xi(\omega)| + \sup_{\omega\in\Omega}|\xi(\omega)-\xi_m(\omega)| < \frac{1}{n}+\frac{1}{m} \le \frac{2}{n}\,.$$

Из этой оценки следует:

$$\forall\, \varepsilon > 0\ \exists\, N \in \mathbb{N}\ :\ \forall\, n, m > N\quad |E\xi_n - E\xi_m| < \varepsilon\,.$$

Таким образом, последовательность $\{E\xi_n\}$ является фундаментальной. По критерию Коши сходимости последовательности в $\mathbb{R}$ из фундаментальности $\{E\xi_n\}$ следует сходимость этой последовательности. ■

Из доказательства предложения 7.1.4 видно, что последовательность дискретных случайных величин ξ_n, равномерно на Ω сходящуюся к случайной величине ξ, можно построить бесконечным множеством способов (достаточно вместо разбиения числовой прямой точками вида $\dfrac{k}{n}$ использовать любое другое разбиение, диаметр которого при $n \to \infty$ стремится к нулю). Поэтому возникает вопрос о независимости определения математического ожидания от выбора последовательности $\{\xi_n\}$.

Предложение 7.1.6. Предел (7.6) не зависит от выбора последовательности дискретных случайных величин $\{\xi_n\}$.

Доказательство. Пусть $\{\xi'_n\}$ и $\{\xi''_n\}$ две последовательности дискретных случайных величин с конечными математическими ожиданиями, равномерно на Ω сходящихся к случайной величине ξ. В силу предложения 7.1.5 пределы $\lim\limits_{n\to\infty} E\xi'_n$ и $\lim\limits_{n\to\infty} E\xi''_n$ существуют и конечны. Предположим, что $\lim\limits_{n\to\infty} E\xi'_n \neq \lim\limits_{n\to\infty} E\xi''_n$. Рассмотрим последовательность дискретных случайных величин, построенную путем «перемешивания» этих двух последовательностей:

$$\xi'_1,\ \xi''_1,\ \xi'_2,\ \xi''_2,\ \xi'_3,\ \xi''_3,\ \ldots\ .$$

Очевидно, она также равномерно на Ω сходится к ξ. Тогда, в силу предложения 7.1.5, предел последовательности математических ожиданий этой последовательности

$$E\xi'_1,\ E\xi''_1,\ E\xi'_2,\ E\xi''_2,\ E\xi'_3,\ E\xi''_3,\ \ldots$$

существует и конечен. Однако это противоречит тому, что у этой последовательности есть две подпоследовательности, имеющие разные пределы. Из полученного противоречия следует равенство

$$\lim_{n\to\infty} E\xi'_n = \lim_{n\to\infty} E\xi''_n. \qquad \blacksquare$$

Таким образом, определение 7.1.2 корректно.

Ⓡ Можно доказать, что свойства (E-I) — (E-V) математического ожидания остаются верными и в общем случае.

Ⓡ Рассмотренное выше двухшаговое определение математического ожидания является частным случаем определения ***интеграла Лебега*** от измеримой функции f, определенной на пространстве с мерой $(\Omega, \mathcal{F}, \mu)$, где Ω — некоторое непустое множество, $\mathcal{F}$ — сигма-алгебра его подмножеств, а μ — конечная сигма-аддитивная мера, определенная на $\mathcal{F}$ (мера называется конечной, если выполнено условие $\mu(\Omega) < \infty$). Интеграл Лебега от функции f по множеству Ω по мере μ обозначают символом $\int\limits_{\Omega} f(\omega)\, d\mu(\omega)$ В контексте теории интеграла Лебега функции, определенные равенством (7.1), называют ***простыми*** измеримыми функциями, а интеграл Лебега для них определяют по формуле (7.2) при условии, что ряд в правой части абсолютно сходится. Далее определение интеграла Лебега распространяется на максимально широкий класс измеримых функций так же, как это сделано для математического ожидания в определении 7.1.2. Таким образом, с использованием интеграла Лебега определение математического ожидания можно дать одним простым равенством:

$$E\xi := \int\limits_{\Omega} \xi(\omega)\, dP(\omega)\,. \tag{7.8}$$

Понятно, что если случайная величина не является дискретной, практическое вычисление математического ожидания по определению 7.1.2, т. е. с помощью формул (7.6) или (7.8), затруднительно. Далее мы рассмотрим формулы, использующиеся для вычисления математического ожидания в важнейшем частном случае, когда недискретная случайная величина абсолютно непрерывна.

Вычисление математического ожидания абсолютно непрерывной случайной величины

Теорема 7.1.7. Пусть ξ — абсолютно непрерывная случайная величина, определенная на вероятностном пространстве $(\Omega, \mathcal{F}, P)$, с плотностью распределения вероятностей $f_\xi(x)$. Тогда

$$E\xi = \int\limits_{-\infty}^{\infty} x f_\xi(x)\, dx\,, \tag{7.9}$$

если интеграл абсолютно сходится.

Доказательство. Действительно, пусть дискретные случайные величины ξ_n определены равенством (7.7). Тогда

$$\begin{aligned} E\xi_n &= \sum_k \frac{k}{n} \cdot P\left(\xi_n = \frac{k}{n}\right) = \\ &= \sum_k \frac{k}{n} \cdot P\left(\frac{k}{n} \le \xi < \frac{k+1}{n}\right) = \sum_k \frac{k}{n} \int\limits_{\frac{k}{n}}^{\frac{k+1}{n}} f_\xi(x)\, dx\,. \end{aligned} \tag{7.10}$$

Предположим, что интеграл (7.9) абсолютно сходится. Тогда абсолютно сходится ряд (7.10). Действительно, в силу того, что на промежутке $\left[\frac{k}{n}; \frac{k+1}{n}\right)$ при $k \ge 0$ верно неравенство $\frac{k}{n} \le |x|$, а при $k < 0$ верно неравенство $-\frac{k+1}{n} \le |x|$, справедлива оценка

$$\sum_{k=-N}^{N} \left| \frac{k}{n} \int\limits_{\frac{k}{n}}^{\frac{k+1}{n}} f_\xi(x)\, dx \right| = \sum_{k=-N}^{-1} \left(-\frac{k}{n}\right) \int\limits_{\frac{k}{n}}^{\frac{k+1}{n}} f_\xi(x)\, dx + \sum_{k=1}^{N} \frac{k}{n} \int\limits_{\frac{k}{n}}^{\frac{k+1}{n}} f_\xi(x)\, dx =$$

$$= \sum_{k=-N}^{-1} \int\limits_{\frac{k}{n}}^{\frac{k+1}{n}} \left(-\frac{k+1}{n}\right) f_\xi(x)\, dx + \sum_{k=1}^{N} \int\limits_{\frac{k}{n}}^{\frac{k+1}{n}} \frac{k}{n} f_\xi(x)\, dx + \sum_{k=-N}^{-1} \int\limits_{\frac{k}{n}}^{\frac{k+1}{n}} \frac{1}{n} f_\xi(x)\, dx =$$

$$\leq \sum_{k=-N}^{-1} \int\limits_{\frac{k}{n}}^{\frac{k+1}{n}} |x| f_\xi(x)\,dx + \sum_{k=1}^{N} \int\limits_{\frac{k}{n}}^{\frac{k+1}{n}} |x| f_\xi(x)\,dx + \frac{1}{n} \sum_{k=-N}^{-1} \int\limits_{\frac{k}{n}}^{\frac{k+1}{n}} f_\xi(x)\,dx =$$

$$= \int\limits_{-N}^{N} |x| f_\xi(x)\,dx + \frac{1}{n} \int\limits_{-N}^{0} f_\xi(x)\,dx \leq \int\limits_{-\infty}^{+\infty} |x| f_\xi(x)\,dx + \frac{1}{n}\,,$$

т. е. частичные суммы ряда из абсолютных величин членов ряда (7.10) ограничены. Это эквивалентно тому, что ряд абсолютно сходится. Докажем, что

$$\lim_{n\to\infty} E\xi_n = \lim_{n\to\infty} \sum_k \frac{k}{n} \int\limits_{\frac{k}{n}}^{\frac{k+1}{n}} f_\xi(x)\,dx = \int\limits_{-\infty}^{\infty} x f_\xi(x)\,dx\,.$$

Действительно,

$$\left| \int\limits_{-\infty}^{\infty} x f_\xi(x)\,dx - \sum_k \frac{k}{n} \int\limits_{\frac{k}{n}}^{\frac{k+1}{n}} f_\xi(x)\,dx \right| = \left| \sum_k \int\limits_{\frac{k}{n}}^{\frac{k+1}{n}} \left(x - \frac{k}{n} \right) f_\xi(x)\,dx \right| \leq$$

$$\leq \sum_k \int\limits_{\frac{k}{n}}^{\frac{k+1}{n}} \left| x - \frac{k}{n} \right| f_\xi(x)\,dx \leq \sum_k \int\limits_{\frac{k}{n}}^{\frac{k+1}{n}} \frac{1}{n} f_\xi(x)\,dx =$$

$$= \frac{1}{n} \sum_k \int\limits_{\frac{k}{n}}^{\frac{k+1}{n}} f_\xi(x)\,dx = \frac{1}{n} \int\limits_{-\infty}^{\infty} f_\xi(x)\,dx = \frac{1}{n} \xrightarrow[n\to\infty]{} 0\,. \qquad \blacksquare$$

R Можно доказать, что математическое ожидание абсолютно непрерывной случайной величины ξ существует тогда и только тогда, когда интеграл (7.9) абсолютно сходится, т. е.

$$\int\limits_{-\infty}^{+\infty} |x| f_\xi(x)\,dx < \infty\,.$$

Если этот интеграл сходится, он равен $E|\xi|$, поэтому утверждение «математическое ожидание ξ существует» часто записывают в виде условия $E|\xi| < \infty$.

Справедливо обобщение формулы (7.9), которое мы примем без доказательства.

Следствие 7.1.8. Пусть ξ — абсолютно непрерывная случайная величина, определенная на вероятностном пространстве $(\Omega, \mathcal{F}, P)$, $f_\xi(x)$ — ее плотность распределения вероятностей, $g : \mathbb{R} \to \mathbb{R}$ — функция, измеримая по Борелю[a]. Тогда

$$Eg(\xi) = \int\limits_{-\infty}^{\infty} g(x) f_\xi(x)\, dx\,. \tag{7.11}$$

[a]Функция $g : \mathbb{R} \to \mathbb{R}$ называется измеримой по Борелю, если она $\mathcal{B}(\mathbb{R})|\mathcal{B}(\mathbb{R})$-измерима (см. определение 5.1.1). Подавляющее большинство функций, с которыми приходится иметь дело в различных приложениях, обладают этим свойством. Например, любая непрерывная функция, любая монотонная функция измеримы по Борелю. Все элементарные функции измеримы по Борелю.

Следующее следствие дает обобщение формулы (7.5) на случай абсолютно непрерывных случайных величин.

Следствие 7.1.9. Пусть $\vec{\xi} = (\xi_1, \xi_2, \ldots, \xi_n)^T$ — случайный вектор с абсолютно непрерывными компонентами, распределение вероятностей которого задано плотностью $f_{\vec{\xi}}(\overline{x}) = f_{\xi_1,\xi_2,\ldots,\xi_n}(x_1, x_2, \ldots, x_n)$, а $g : \mathbb{R}^n \to \mathbb{R}$ — функция, измеримая по Борелю[a]. Тогда

$$Eg(\xi) = \int \cdots \int\limits_{\mathbb{R}^n} g(\overline{x}) f_{\vec{\xi}}(\overline{x})\, d\overline{x}\,. \tag{7.12}$$

[a]Функция $g : \mathbb{R}^n \to \mathbb{R}$ называется измеримой по Борелю, если она $\mathcal{B}(\mathbb{R}^n)|\mathcal{B}(\mathbb{R})$-измерима.

Рассмотрим примеры вычисления математического ожидания абсолютно непрерывных случайных величин, имеющих распределения вероятностей, введенные в главе 5.

■ Пример 7.1.6. Пусть $\xi \sim U(a, b)$, тогда $E\xi = \dfrac{a+b}{2}$.

Действительно, по формуле (7.9) получаем

$$E\xi = \int\limits_{-\infty}^{\infty} x f_\xi(x)\, dx = \frac{1}{b-a} \int\limits_a^b x\, dx = \frac{x^2}{2(b-a)}\bigg|_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}\,. \quad ■$$

■ Пример 7.1.7. Пусть $\xi \sim \mathrm{Exp}(\lambda)$, тогда $E\xi = \dfrac{1}{\lambda}$.

Также, используя формулу (7.9), получаем

$$E\xi = \int\limits_{-\infty}^{\infty} x f_\xi(x)\,dx = \lambda \int\limits_{0}^{\infty} x e^{-\lambda x}\,dx = -\int\limits_{0}^{\infty} x\,de^{-\lambda x} =$$

$$= -xe^{-\lambda x}\Big|_0^\infty + \int\limits_0^\infty e^{-\lambda x}\,dx = 0 - \frac{e^{-\lambda x}}{\lambda}\Big|_0^\infty = \frac{1}{\lambda}\,. \qquad \blacksquare$$

▪ Пример 7.1.8. Пусть $\xi \sim N(a, \sigma^2)$, тогда $E\xi = a$.

Пусть сначала $\xi \sim N(0, 1)$, тогда по формуле (7.9) получаем:

$$E\xi = \int\limits_{-\infty}^{\infty} x f_\xi(x)\,dx = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} x e^{-\frac{x^2}{2}}\,dx = 0\,,$$

так как под интегралом нечетная функция. Далее, пусть $\xi \sim N(a, \sigma^2)$, тогда, пользуясь предложением 5.5.3, представим ξ в виде $\xi = \sigma\eta + a$, где $\eta \sim N(0, 1)$, и с помощью свойств (E-I) и (E-II) математического ожидания получим

$$E\xi = E(\sigma\eta + a) = \sigma E\eta + Ea = a\,. \qquad \blacksquare$$

Рассмотрим следующее важное свойство математического ожидания, связанное с понятием независимости случайных величин.

Теорема 7.1.10. Если случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ независимы, то

$$E\prod_{i=1}^{n} \xi_i = \prod_{i=1}^{n} E\xi_i\,. \tag{7.13}$$

Доказательство. Ограничимся проверкой равенства (7.13) в двух наиболее распространенных в приложениях частных случаях.

а) Пусть $\xi_1, \xi_2, \ldots, \xi_n$ — дискретные случайные величины, их совместное распределение задано вероятностями

$$p_{j_1, j_2, \ldots, j_n} = P\big(\xi_1 = x^1_{j_1}, \xi_2 = x^2_{j_2}, \ldots, \xi_n = x^n_{j_n}\big)\,,$$

а маргинальные распределения вероятностей заданы вероятностями

$$P(\xi = x^i_j) = p^i_j\,.$$

Тогда, пользуясь формулой (7.5), а затем критерием независимости дискретных случайных величин (Теорема 6.5.1), получаем:

$$E\prod_{k=1}^{n}\xi_k = \sum_{j_1,j_2,\ldots,j_n} x^1_{j_1}x^2_{j_2}\ldots x^n_{j_n}p_{j_1,j_2,\ldots,j_n} = \sum_{j_1,j_2,\ldots,j_n} x^1_{j_1}x^2_{j_2}\ldots x^n_{j_n}\prod_{i=1}^{n}p^i_{j_i} =$$

$$= \sum_{j_1,j_2,\ldots,j_n}\prod_{i=1}^{n}x^i_{j_i}p^i_{j_i} = \prod_{i=1}^{n}\sum_{j_i}x^i_{j_i}p^i_{j_i} = \prod_{i=1}^{n}E\xi_i\,.$$

b) Пусть случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ абсолютно непрерывные, их совместное распределение задано плотностью совместного распределения $f_{\vec{\xi}} = f_{\xi_1,\ldots,\xi_n}$, а $f_{\xi_i}, i = 1,\ldots,n$ — плотности маргинальных распределений вероятностей. Тогда, пользуясь формулой (7.12), а затем критерием независимости абсолютно непрерывных случайных величин (Теорема 6.5.3), получаем:

$$E\prod_{k=1}^{n}\xi_k = \underset{\mathbb{R}^n}{\int\cdots\int}\prod_{i=1}^{n}x_i\cdot f_{\vec{\xi}}(x_1,x_2,\ldots,x_n)\,dx_1dx_2\ldots dx_n =$$

$$= \underset{\mathbb{R}^n}{\int\cdots\int}\prod_{i=1}^{n}x_i\prod_{i=1}^{n}f_{\xi_i}(x_i)\,dx_1dx_2\ldots dx_n =$$

$$= \underset{\mathbb{R}^n}{\int\cdots\int}\prod_{i=1}^{n}\left[x_if_{\xi_i}(x_i)\right]\,dx_1dx_2\ldots dx_n =$$

$$= \prod_{i=1}^{n}\int_{\mathbb{R}}x_if_{\xi_i}(x_i)\,dx_i = \prod_{i=1}^{n}E\xi_i\,. \qquad \blacksquare$$

Определив математическое ожидание (момент первого порядка) случайной величины, мы можем с его помощью определить моменты высших порядков.

7.2. Моменты высших порядков

Определение 7.2.1. ***Моментом k-го порядка случайной величины ξ относительно a*** называют величину $E(\xi - a)^k$. При $a = 0$ ее называют просто ***моментом k-го порядка***, а при $a = E\xi$ — ***центральным моментом k-го порядка***. Величину $E|\xi - a|^k$ называют ***абсолютным моментом k-го порядка относительно*** a. При $a = 0$ эту величину называют просто ***абсолютным моментом k-го порядка***, а при $a = E\xi$ — ***абсолютным центральным моментом k-го порядка***.

Среди всех моментов случайной величины чаще всего используется центральный момент второго порядка.

Определение 7.2.2. Центральный момент второго порядка случайной величины ξ называют ***дисперсией*** и обозначают символом $D\xi$. Таким образом,

$$D\xi := E(\xi - E\xi)^2 .$$

Определенная как среднее значение квадрата отклонения случайной величины от ее математического ожидания, дисперсия характеризует возможный разброс ее значений.

Предложение 7.2.1. Дисперсия обладает следующими свойствами:
(D-I) $\forall\ a, b, \in \mathbb{R}\ \ D(a\xi + b) = a^2 D\xi$;
(D-II) $D\xi = E\xi^2 - (E\xi)^2$;
(D-III) $\min\limits_{a\in\mathbb{R}} E(\xi - a)^2 = E(\xi - E\xi)^2 = D\xi$;

Доказательство. Пользуясь линейностью математического ожидания, получаем:

$$D(a\xi + b) = E\left(a\xi + b - E(a\xi + b)\right)^2 = E\left(a\xi + b - aE\xi - b\right)^2 = \\ = E\left(a(\xi - E\xi)\right)^2 = a^2 E(\xi - E\xi)^2 = a^2 D\xi .$$

Аналогично,

$$D\xi = E(\xi - E\xi)^2 = E\left(\xi^2 - 2\xi E\xi + (E\xi)^2\right) = \\ = E\xi^2 - 2E\xi \cdot E\xi + (E\xi)^2 = E\xi^2 - (E\xi)^2 .$$

Свойства (D-I) и (D-II) доказаны. Для того чтобы доказать свойство (D-III), запишем

$$E(\xi - a)^2 = E\left((\xi - E\xi) - (a - E\xi)\right)^2 = \\ = E\left((\xi - E\xi)^2 - 2(\xi - E\xi)(a - E\xi) + (a - E\xi)^2\right) = \\ = D\xi - 2(a - E\xi) \cdot E(\xi - E\xi) + (a - E\xi)^2 = D\xi + (a - E\xi)^2 ,$$

поскольку $E(\xi - E\xi) = E\xi - E\xi = 0$. Из полученного представления видно, что минимум правой части равенства достигается при $a = E\xi$ и равен $D\xi$. ■

Для практического вычисления дисперсии удобно использовать свойство (D-II) в сочетании с формулами (7.4) и (7.11).

■ Пример 7.2.1. Пусть $\xi \sim B(n,p)$, тогда $D\xi = npq$, где $q = 1-p$.

Действительно, найдем момент второго порядка ξ:

$$E\xi^2 = \sum_{k=0}^{n} k^2 C_n^k p^k q^{n-k} = \sum_{k=0}^{n} (k^2 - k) C_n^k p^k q^{n-k} + E\xi =$$
$$= p^2 \sum_{k=0}^{n} k(k-1) C_n^k p^{k-2} q^{n-k} + E\xi = p^2 \frac{d^2}{dp^2}(p+q)^n + E\xi =$$
$$= p^2 n(n-1) + np = p^2 n^2 + pn(1-p) = (np)^2 + npq\,.$$

В результате получаем:

$$D\xi = E\xi^2 - (E\xi)^2 = npq\,.$$ ■

■ Пример 7.2.2. Пусть $\xi \sim N(a, \sigma^2)$. Тогда $D\xi = \sigma^2$.

Действительно, пусть сначала $\xi \sim N(0,1)$. Тогда

$$D\xi = E\xi^2 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-\frac{x^2}{2}}\, dx = -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x\, de^{-\frac{x^2}{2}} =$$
$$= -\frac{1}{\sqrt{2\pi}} x e^{-\frac{x^2}{2}} \Big|_{-\infty}^{+\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\, dx = 1\,.$$

(R) Еще один способ получить этот результат — использовать гамма-функцию Эйлера и ее свойства (см. Приложение B):

$$E\xi^2 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-\frac{x^2}{2}}\, dx = \sqrt{\frac{2}{\pi}} \int_{0}^{+\infty} x^2 e^{-\frac{x^2}{2}}\, dx = \sqrt{\frac{2}{\pi}} \int_{0}^{+\infty} x e^{-\frac{x^2}{2}}\, d\left(\frac{x^2}{2}\right) =$$
$$= \frac{2}{\sqrt{\pi}} \int_{0}^{+\infty} \sqrt{s} e^{-s} ds = \frac{2}{\sqrt{\pi}} \Gamma\left(\frac{3}{2}\right) = \frac{1}{\sqrt{\pi}} \Gamma\left(\frac{1}{2}\right) = 1\,.$$

Далее, пусть $\xi \sim N(a, \sigma^2)$, тогда (Предложение 5.5.3) $\xi = \sigma\eta + a$, где $\eta \sim N(0,1)$. Тогда по свойству (D-I) получаем $D\xi = \sigma^2 D\eta = \sigma^2$. ∎

Следующая теорема и следствие из нее связывают между собой моментные и вероятностные характеристики случайных величин.

Теорема 7.2.2. — Неравенство Чебышева — I. Пусть ξ — случайная величина, определенная на вероятностном пространстве $(\Omega, \mathcal{F}, P)$, $\xi \geq 0$. Тогда для любого $k \in \mathbb{N}$, если $E\xi^k < \infty$, то

$$\forall\, \varepsilon > 0 \;\; P\left(\xi > \varepsilon\right) \leq \frac{E\xi^k}{\varepsilon^k}\,. \tag{7.14}$$

Доказательство. Пусть $\varepsilon > 0$. Определим вспомогательную случайную величину η, положив

$$\eta = \begin{cases} \varepsilon\,, & \xi > \varepsilon\,, \\ 0\,, & \xi \leq \varepsilon\,. \end{cases}$$

Тогда $\eta \leq \xi$, а значит для любого $k \in \mathbb{N}$ $\;\eta^k \leq \xi^k$. Пользуясь свойством (E-IV) математического ожидания, получаем $E\eta^k \leq E\xi^k$, но по определению математического ожидания дискретной случайной величины $E\eta^k = \varepsilon^k P\left(\xi > \varepsilon\right)$, в результате получаем неравенство (7.14). ∎

Применяя неравенство (7.14) к неотрицательной случайной величине $|\xi - E\xi|$ при $k = 2$, получаем следствие, которое также называют неравенством Чебышева.

Следствие 7.2.3. — Неравенство Чебышева — II. Пусть ξ — случайная величина, определенная на вероятностном пространстве $(\Omega, \mathcal{F}, P)$, $D\xi < \infty$. Тогда

$$\forall\, \varepsilon > 0 \;\; P\big(|\xi - E\xi| > \varepsilon\big) \leq \frac{D\xi}{\varepsilon^2}\,. \tag{7.15}$$

Неравенство (7.15) показывает, что дисперсия характеризует способность случайной величины отклоняться от своего математического ожидания: чем меньше дисперсия, тем менее вероятны большие отклонения.

Размерность дисперсии равна квадрату размерности случайной величины, поэтому для характеристики ее способности отклоняться от среднего значения удобнее использовать корень из дисперсии.

Определение 7.2.3. ***Стандартное отклонение*** σ_ξ случайной величины ξ определяется равенством

$$\sigma_\xi := \sqrt{D\xi}\,.$$

Из неравенства Чебышева (7.15) вытекает правило трех сигм.

Правило трех сигм: С вероятностью, не меньшей $\dfrac{8}{9}$, любая случайная величина ξ с $D\xi < \infty$ отклоняется от своего математического ожидания не более, чем на $3\sigma_\xi$.

Действительно, полагая в неравенстве (7.15) $\varepsilon = 3\sigma_\xi$, получаем

$$P\left(|\xi - E\xi| \le 3\sigma_\xi\right) = 1 - P\left(|\xi - E\xi| > 3\sigma_\xi\right) \ge 1 - \frac{D\xi}{9\sigma_\xi^2} = \frac{8}{9}\,.$$

R Правило трех сигм выводится из неравенства Чебышева, представляющего собой довольно грубую оценку вероятности $P\left(|\xi - E\xi| > \varepsilon\right)$. Для многих классов случайных величин вероятность $P\left(|\xi - E\xi| \le 3\sigma_\xi\right)$ оказывается значительно более близкой к единице. Например, для гауссовских случайных величин она равна приблизительно $0,9973$.

Отметим еще одно свойство дисперсии, которое вытекает из неравенства Чебышева.

Предложение 7.2.4. **(D-IV)** $D\xi = 0 \;\Leftrightarrow\; \xi = E\xi$ с вероятностью 1.

Иначе говоря, равенство нулю дисперсии случайной величины эквивалентно тому, что она с вероятностью 1 является константой.

Доказательство. Если $\xi = E\xi = \mathrm{Const}$ с вероятностью 1, по свойству (D-I), получаем $D\xi = 0$. Обратно, пусть $D\xi = 0$. Рассмотрим событие $A = \{\xi = E\xi\}$ и последовательность событий $A_n = \left\{|\xi - E\xi| \le \dfrac{1}{n}\right\}$, $n \in \mathbb{N}$. Поскольку для любого $n \in \mathbb{N}$ имеют место вложения $A_n \supseteq A_{n+1}$ и $A = \bigcap_{n=1}^{\infty} A_n$, то по свойству непрерывности вероятности (P-9) имеем $P(A) = \lim_{n\to\infty} P(A_n)$, при этом

$$P(A_n) = 1 - P\left(|\xi - E\xi| > \frac{1}{n}\right) \geq 1 - \frac{D\xi}{1/n^2} = 1\,,$$

следовательно $P(A) = 1$. ■

Аналогично доказывается следующее предложение.

Предложение 7.2.5. $E\xi^2 = 0 \;\Leftrightarrow\; \xi = 0$ с вероятностью 1.

R Существуют случайные величины, у которых определены не все моменты, и даже такие, у которых все моменты не определены. Например, для случайной величины ξ, распределенной по закону Коши, не существуют ни $E\xi$, ни другие моменты. Действительно, пусть $\xi \sim C(x_0, \gamma)$, тогда $E|\xi^k| = \infty$ для любого $k \in \mathbb{N}$. Действительно,

$$E\xi^k = \frac{\gamma}{\pi}\int_{-\infty}^{+\infty} \frac{x^k}{\gamma^2 + (x - x_0)^2}\,dx\,,$$

где в правой части равенства расходящийся несобственный интеграл.

R Совокупность определенных на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ случайных величин ξ, имеющих конечный момент второго порядка ($E|\xi|^2 < \infty$), является линейным пространством над полем $\mathbb{R}$. При этом величина

$$\|\xi\| := \sqrt{E|\xi|^2} \tag{7.16}$$

обладает всеми свойствами нормы, если договориться отождествлять элементы ξ и η этого пространства, удовлетворяющие условию $E|\xi - \eta|^2 = 0$ (такое допущение вполне естественно, так как в силу предложения 7.2.5 $E|\xi - \eta|^2 = 0 \;\Leftrightarrow\; \xi = \eta$ с вероятностью 1).

Таким образом, линейное пространство определенных на $(\Omega, \mathcal{F}, P)$ случайных величин с конечным вторым моментом, оснащенное нормой, заданной равенством (7.16), является нормированным пространством. Его обозначают символом $L_2(\Omega, \mathcal{F}, P)$ или кратко $L_2(\Omega)$. Таким образом, величина $\|\xi - \eta\| = \sqrt{E(\xi - \eta)^2}$ играет роль расстояния между случайными величинами в этом пространстве. Это расстояние называют ***среднеквадратическим***.

7.3. Ковариационный момент. Коэффициент корреляции

Определение 7.3.1. Пусть случайные величины ξ и η заданы на одном вероятностном пространстве $(\Omega, \mathcal{F}, P)$. ***Ковариационным моментом*** или ***ковариацией*** случайных величин ξ и η называется величина

$$\mathrm{Cov}\,(\xi, \eta) := E\big((\xi - E\xi)(\eta - E\eta)\big)\,.$$

Ковариационный момент возникает при вычислении дисперсии суммы случайных величин.

Предложение 7.3.1. **(D-V)** $D(\xi + \eta) = D\xi + D\eta + 2\mathrm{Cov}\,(\xi, \eta)\,.$

Доказательство.

$$D(\xi + \eta) = E\big((\xi + \eta) - E(\xi + \eta)\big)^2 = E\big((\xi - E\xi) + (\eta - E\eta)\big)^2 =$$
$$= E\big((\xi - E\xi)^2 + (\eta - E\eta)^2 + 2(\xi - E\xi)(\eta - E\eta)\big) =$$
$$= D\xi + D\eta + 2\mathrm{Cov}\,(\xi, \eta)\,.$$ ∎

Рассмотрим свойства ковариации.

Предложение 7.3.2. Для любых случайных величин ξ и η и констант α и $\beta \in \mathbb{R}$

(C-I) $\mathrm{Cov}\,(\alpha\xi_1 + \beta\xi_2, \eta) = \alpha\mathrm{Cov}\,(\xi_1, \eta) + \beta\mathrm{Cov}\,(\xi_2, \eta)$;

(C-II) $\mathrm{Cov}\,(\alpha\xi + \beta, \eta) = \alpha\mathrm{Cov}\,(\xi, \eta)$;

(C-III) $\mathrm{Cov}\,(\xi, \eta) = \mathrm{Cov}\,(\eta, \xi)$;

(C-IV) $\mathrm{Cov}\,(\xi, \xi) = D\xi$;

(C-V) $\mathrm{Cov}\,(\xi, \eta) = E(\xi\eta) - E\xi \cdot E\eta$;

(C-VI) Если ξ и η независимы, то $\mathrm{Cov}\,(\xi, \eta) = 0\,.$

Докажите это предложение самостоятельно, используя определение ковариации и свойства математического ожидания.

R Утверждение, обратное к (C-VI) неверно. Действительно, пусть $\xi \sim U\{-1; 0; 1\}$, $\eta = |\xi|$. Тогда $E\xi = 0\,, E\eta = 2/3\,, \xi\eta = \xi$, следовательно, $E(\xi\eta) = 0$. В результате, используя свойство (C-V), получаем

$\mathrm{Cov}\,(\xi, \eta) = 0$. При этом ξ и η не являются независимыми, так как η представляет собой функцию от ξ и ее значение полностью определяется тем, какое значение примет ξ.

Отметим еще одно свойство дисперсии, которое вытекает из свойств (D-V) и (C-VI).

Следствие 7.3.3. **(D-VI)** Если случайные величины ξ и η независимы, то $D(\xi + \eta) = D\xi + D\eta$.

Из свойства (C-VI) следует, что если ковариация двух случайных величин отлична от нуля, то эти случайные величины зависимы. Следующее понятие используется для описания зависимостей между компонентами случайного вектора.

Определение 7.3.2. Пусть $\vec{\xi} = (\xi_1, \xi_2, \ldots, \xi_n)^T$ — случайный вектор. ***Ковариационной матрицей*** случайного вектора $\vec{\xi}$ называется матрица $\Lambda = \big(\lambda_{ij}\big)_{n\times n}$, где $\lambda_{ij} = \mathrm{Cov}\,(\xi_i, \xi_j)$.

Предложение 7.3.4. Ковариационная матрица случайного вектора $\vec{\xi} = (\xi_1, \xi_2, \ldots, \xi_n)^T$ обладает следующими свойствами:

(Λ-I) Ковариационная матрица симметрическая: $\Lambda^T = \Lambda$;

(Λ-II) $\forall\, \overline{x} \in \mathbb{R}^n \ \ \big(\Lambda\overline{x}, \overline{x}\big) = D\left(\sum\limits_{i=1}^{n} x_i\xi_i\right)$;

(Λ-III) Ковариационная матрица является неотрицательно определенной: $\forall\, \overline{x} \in \mathbb{R}^n \ \ \big(\Lambda\overline{x}, \overline{x}\big) \geq 0$;

(Λ-IV) $\exists\, \overline{x} \neq \overline{0} \ : \ \big(\Lambda\overline{x}, \overline{x}\big) = 0 \quad \Leftrightarrow \quad \exists\, \overline{x} = (x_1, \ldots, x_n)^T \neq \overline{0} \ :$ $\sum\limits_{i=1}^{n} x_i\xi_i = \sum\limits_{i=1}^{n} x_i E\xi_i$ с вероятностью 1.

Доказательство. Симметричность матрицы Λ следует из свойства (C-III):

$$\lambda_{ij} = \mathrm{Cov}\,(\xi_i, \xi_j) = \mathrm{Cov}\,(\xi_j, \xi_i) = \lambda_{ji}\,.$$

Далее, для любого $\overline{x} \in \mathbb{R}^n$ имеем:

$$\begin{aligned}\big(\Lambda\overline{x}, \overline{x}\big) &= \sum_{i,j=1}^{n} \lambda_{ij} x_i x_j = \sum_{i,j=1}^{n} \mathrm{Cov}\,(\xi_i, \xi_j) x_i x_j = \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} \mathrm{Cov}\,(x_i\xi_i, x_j\xi_j) = \mathrm{Cov}\,\left(\sum_{i=1}^{n} x_i\xi_i, \sum_{j=1}^{n} x_j\xi_j\right) = \\ &= D\left(\sum_{i=1}^{n} x_i\xi_i\right) \geq 0\,.\end{aligned}$$

Таким образом, доказано свойство (Λ-II). Свойство (Λ-III) очевидно следует из (Λ-II). Свойство (Λ-IV) вытекает из свойства дисперсии (D-IV) и свойства (Λ-II). ■

R Свойство (Λ-IV) можно сформулировать следующим образом. Ковариационная матрица Λ нестрого определена тогда и только тогда, когда некоторая невырожденная линейная комбинация случайных величин ξ_i является константой с вероятностью 1 (иначе говоря, случайные величины функционально связаны (линейной зависимостью)).

Теорема 7.3.5. — Неравенство Шварца.

$$\big|\mathrm{Cov}\,(\xi, \eta)\big| \leq \sqrt{D\xi} \cdot \sqrt{D\eta}\,. \tag{7.17}$$

Доказательство. Пусть ковариационная матрица случайного вектора $\vec{\xi} = (\xi, \eta)^T$ имеет вид

$$\Lambda = \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{pmatrix} = \begin{pmatrix} D\xi & \mathrm{Cov}\,(\xi, \eta) \\ \mathrm{Cov}\,(\xi, \eta) & D\eta \end{pmatrix}.$$

Пользуясь неотрицательной определенностью Λ и тем, что $\lambda_{11} \geq 0$, получаем:

$$\begin{aligned}&0 \leq \big(\Lambda\overline{x}, \overline{x}\big) = \lambda_{11} x_1^2 + 2\lambda_{12} x_1 x_2 + \lambda_{22} x_2^2 \;\Leftrightarrow \\ &\Leftrightarrow\; D = 4\lambda_{12}^2 - 4\lambda_{11}\lambda_{22} \leq 0 \;\Leftrightarrow\; \lambda_{12}^2 \leq \lambda_{11}\lambda_{22}\,,\end{aligned}$$

откуда следует неравенство (7.17). ■

Для более точного описания степени зависимости случайных величин между собой используют ковариационный момент, нормированный на стандартные отклонения случайных величин.

Определение 7.3.3. ***Коэффициентом корреляции*** случайных величин ξ и η называется величина

$$\rho_{\xi\eta} := \frac{\mathrm{Cov}\,(\xi,\eta)}{\sqrt{D\xi}\cdot\sqrt{D\eta}} = \frac{\mathrm{Cov}\,(\xi,\eta)}{\sigma_\xi\cdot\sigma_\eta}\,. \tag{7.18}$$

В результате, ковариационная матрица, например, двумерного случайного вектора $\vec{\xi} = (\xi,\eta)^T$, может быть записана следующем образом

$$\Lambda = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

где $\sigma_1 = \sigma_\xi\,, \sigma_2 = \sigma_\eta\,, \rho = \rho_{\xi\eta}$.

Предложение 7.3.6. Коэффициент корреляции обладает следующими свойствами:
(ρ-I) $\rho_{\xi\eta} \in [-1;1]$;
(ρ-II) Если случайные величины ξ и η независимы, то $\rho_{\xi\eta} = 0$;
(ρ-III) $|\rho_{\xi\eta}| = 1 \;\Leftrightarrow\; \exists\, a,b \in \mathbb{R} \;:\; \xi = a\eta + b = E\xi \pm \dfrac{\sigma_1}{\sigma_2}(\eta - E\eta)$.

Доказательство. Свойство (ρ-I) сразу следует из неравенства Шварца. Свойство (ρ-II) — следствие свойства (C-VI) ковариационного момента. Докажем свойство (ρ-III). Пусть $\xi = a\eta + b$, тогда

$$\mathrm{Cov}\,(\xi,\eta) = \mathrm{Cov}\,(a\eta+b,\eta) = a\mathrm{Cov}\,(\eta,\eta) = aD\eta = a\sigma_\eta^2\,,$$
$$D\xi = D(a\eta+b) = a^2D\eta = a^2\sigma_\eta^2\,,$$

следовательно, $\rho_{\xi\eta} = \dfrac{a\sigma_\eta^2}{|a|\sigma_\eta^2} = \mathrm{sgn}\,a\,.$

Обратно, пусть $|\rho_{\xi\eta}| = 1$, тогда ковариационная матрица случайного вектора $\vec{\xi} = (\xi,\eta)^T$ имеет вид

$$\Lambda = \begin{pmatrix} \sigma_1^2 & \pm\sigma_1\sigma_2 \\ \pm\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

Пользуясь свойством (Λ-II), для любого $\overline{x} = (x_1,x_2)^T \in \mathbb{R}^2$ получаем:

$$D(x_1\xi + x_2\eta) = (\Lambda\overline{x},\overline{x}) = \sigma_1^2x_1^2 \pm 2\sigma_1\sigma_2x_1x_2 + \sigma_2^2x_2^2 = (x_1\sigma_1 \pm x_2\sigma_2)^2\,.$$

Выберем x_1 и x_2 так, чтобы выполнялось равенство $x_1\sigma_1 \pm x_2\sigma_2 = 0$, например, $x_1 = 1\,, x_2 = \mp\dfrac{\sigma_1}{\sigma_2}$. Тогда, в силу свойства (D-IV), с вероятностью 1 будет выполняться равенство $x_1\xi \pm x_2\eta = x_1E\xi \pm x_2E\eta$, откуда получаем

$$\xi = E\xi \pm \frac{\sigma_1}{\sigma_2}(\eta - E\eta) = a\eta + b\,,$$

где $a = \pm\dfrac{\sigma_1}{\sigma_2}\,, b = E\xi \mp \dfrac{\sigma_1}{\sigma_2}E\eta$. ■

R Если коэффициент корреляции случайных величин равен нулю, их называют ***некоррелированными***. Свойство (ρ-II) означает, что независимость случайных величин влечет за собой их некоррелированность. Обратное, вообще говоря, неверно. Это следует из замечания на стр. 120 о свойстве (C-VI) ковариационного момента, поскольку $\rho_{\xi\eta} = 0 \;\Leftrightarrow\; \mathrm{Cov}\,(\xi, \eta) = 0$.

Из свойства (ρ-II) следует, что ненулевой коэффициент корреляции означает наличие зависимости между случайными величинами. При этом абсолютная величина коэффициента корреляции служит мерой зависимости между ними. Чем она больше, тем степень зависимости больше. Максимальное по абсолютной величине значение $\rho_{\xi\eta}$, равное 1 или -1, означает наличие жесткой функциональной (линейной) зависимости одной случайной величины от другой. При этом из доказательства свойства (ρ-III) видно, что знак коэффициента корреляции совпадает со знаком углового коэффициента этой линейной зависимости.

7.4. Линейная среднеквадратическая регрессия

Задача линейной среднеквадратической регрессии одной случайной величины на другую относится к задачам прогнозирования. Формулируется она следующим образом. Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, P)$, соответствующем некоторому случайному эксперименту, определены случайные величины ξ и η (или случайный вектор $\vec{\xi} = (\xi, \eta)^T$) с конечными моментами второго порядка. Известно, что они зависимы ($\rho_{\xi\eta} \neq 0$) и у нас есть возможность наблюдать, какое значение в результате проведения случайного эксперимента принимает ξ. Будем строить прогноз ожидаемого значения η с помощью линейной функции $y = ax+b$, при этом качество прогноза будем оценивать по норме пространства $L_2(\Omega)$, т. е. в среднеквадратичном (см. замечание на стр. 119).

Таким образом, мы приходим к следующей задаче.

Найти линейную функцию $y = ax + b$ такую, что

$$E\big(\eta - (a\xi + b)\big)^2 = \min_{\alpha,\beta\in\mathbb{R}} E\big(\eta - (\alpha\xi + \beta)\big)^2 .$$

Для того чтобы решить эту задачу, преобразуем величину, которую требуется минимизировать:

$$E\big(\eta - (\alpha\xi + \beta)\big)^2 = E\big(\eta - E\eta - \alpha(\xi - E\xi) + E\eta - \alpha E\xi - \beta\big)^2 .$$

Вводя обозначение $\gamma = E\eta - \alpha E\xi - \beta$, возводя в квадрат сумму $(\eta - E\eta) - \alpha(\xi - E\xi) + \gamma$ и пользуясь свойствами математического ожидания, дисперсии и ковариации, получим

$$\begin{aligned} E\big(\eta - (\alpha\xi + \beta)\big)^2 &= E\big((\eta - E\eta)^2 + \alpha^2(\xi - E\xi)^2 + \gamma^2 + \\ &\quad + 2\gamma(\eta - E\eta) - 2\gamma\alpha(\xi - E\xi) - 2\alpha(\eta - E\eta)(\xi - E\xi)\big) = \\ &= D\eta + \alpha^2 D\xi + \gamma^2 + 0 + 0 - 2\alpha \mathrm{Cov}\,(\xi,\eta) = \\ &= \sigma_\eta^2 + \alpha^2\sigma_\xi^2 + \gamma^2 - 2\alpha\rho_{\xi\eta}\sigma_\eta\sigma_\xi = \\ &= \sigma_\eta^2 + \gamma^2 + \sigma_\xi^2\left(\alpha^2 - 2\alpha\rho_{\xi\eta}\frac{\sigma_\eta}{\sigma_\xi}\right) = \\ &= \sigma_\eta^2 + \gamma^2 + \sigma_\xi^2\left(\alpha^2 - 2\alpha\rho_{\xi\eta}\frac{\sigma_\eta}{\sigma_\xi} + \rho_{\xi\eta}^2\frac{\sigma_\eta^2}{\sigma_\xi^2}\right) - \rho_{\xi\eta}^2\sigma_\eta^2 = \\ &= \sigma_\eta^2(1 - \rho_{\xi\eta}^2) + (E\eta - \alpha E\xi - \beta)^2 + \sigma_\xi^2\left(\alpha - \rho_{\xi\eta}\frac{\sigma_\eta}{\sigma_\xi}\right)^2 . \end{aligned}$$

Из полученного представления видно, что минимум $E\big(\eta - (\alpha\xi + \beta)\big)^2$ равен $\sigma_\eta^2(1 - \rho_{\xi\eta}^2)$ и достигается при выполнении условий

$$\begin{cases} E\eta - \alpha E\xi - \beta = 0\,, \\ \alpha - \rho_{\xi\eta}\dfrac{\sigma_\eta}{\sigma_\xi} = 0\,. \end{cases}$$

Из этой системы уравнений находим $\alpha = \rho_{\xi\eta}\dfrac{\sigma_\eta}{\sigma_\xi}$, $\beta = E\eta - \rho_{\xi\eta}\dfrac{\sigma_\eta}{\sigma_\xi}E\xi$. Таким образом, решением задачи линейной среднеквадратической регрессии случайной величины η на случайную величину ξ является линейная функция

$$y = ax + b = E\eta + \rho_{\xi\eta}\frac{\sigma_\eta}{\sigma_\xi}(x - E\xi)\,, \tag{7.19}$$

а наилучший в среднеквадратическом линейный прогноз значения η по наблюдаемому значению ξ дает случайная величина

$$E\eta + \rho_{\xi\eta}\frac{\sigma_\eta}{\sigma_\xi}(\xi - E\xi)\,.$$

При этом

$$\min_{\alpha,\beta\in\mathbb{R}} E\big(\eta - (\alpha\xi + \beta)\big)^2 = \sigma_\eta^2(1 - \rho_{\xi\eta}^2)\,.$$

Определение 7.4.1. Прямая, заданная уравнением (7.19), называется ***прямой линейной среднеквадратической регрессии $\boldsymbol{\eta}$ на $\boldsymbol{\xi}$***. Меняя ролями переменные, получаем уравнение ***прямой линейной среднеквадратической регрессии $\boldsymbol{\xi}$ на $\boldsymbol{\eta}$***:

$$x = ay + b = E\xi + \rho_{\xi\eta}\frac{\sigma_\xi}{\sigma_\eta}(y - E\eta)\,. \tag{7.20}$$

R Полученный результат согласуется со свойством (ρ-III) коэффициента корреляции. Действительно, при $|\rho_{\xi\eta}| = 1$ $\min\limits_{\alpha,\beta\in\mathbb{R}} E\big(\xi - (\alpha\eta + \beta)\big)^2 = 0$ и достигается он при $a\eta + b = E\xi \pm \dfrac{\sigma_1}{\sigma_2}(\eta - E\eta)$, так что в результате имеет место равенство

$$\xi = E\xi \pm \frac{\sigma_1}{\sigma_2}(\eta - E\eta)\,.$$

Задача 7.4.1. Совместное распределение вероятностей случайных величин ξ и η задано плотностью

$$f_{\xi\eta}(x, y) = a(1 - x + y)\cdot\mathbf{1}_{[0;1]\times[0;1]}(x, y)\,.$$

Найти значение параметра a, $E\xi$, $E\eta$, $D\xi$, $D\eta$, $\mathrm{Cov}\,(\xi,\eta)$, $\rho_{\xi\eta}$ и построить прямые линейной среднеквадратической регрессии ξ на

η и η на ξ.

Решение. Для нахождения параметра a воспользуемся свойством (ff-II) плотности совместного распределения.

$$\iint\limits_{\mathbb{R}^2} f_{\xi\eta}(x,y)\,dx\,dy = a\int_0^1 dx\int_0^1 (1-x+y)\,dy =$$
$$= a\int_0^1 \left(1 - x + \frac{y^2}{2}\Big|_0^1\right)\,dx = a\int_0^1 \left(\frac{3}{2} - x\right)\,dx =$$
$$= a\left(\frac{3}{2} - \frac{x^2}{2}\right)\Big|_0^1 = a\,.$$

В силу свойства (ff-II) $a = 1$.

Найдем плотности маргинальных распределений вероятностей. Пользуясь формулой (6.6), получим:

$$f_\xi(x) = \int\limits_{-\infty}^{+\infty} f_{\xi\eta}(x,y)\,dy = \begin{cases} 0\,, & x \notin [0;1]\,,\\ \int\limits_0^1 (1-x+y)\,dy = \dfrac{3}{2} - x\,, & x \in [0;1]\,. \end{cases}$$

$$f_\eta(y) = \int\limits_{-\infty}^{+\infty} f_{\xi\eta}(x,y)\,dx = \begin{cases} 0\,, & y \notin [0;1]\,,\\ \int\limits_0^1 (1-x+y)\,dx = \dfrac{1}{2} + y\,, & y \in [0;1]\,. \end{cases}$$

Таким образом,

$$f_\xi(x) = \left(\frac{3}{2} - x\right)\cdot \mathbf{1}_{[0;1]}(x)\,, f_\eta(x) = \left(\frac{1}{2} + x\right)\cdot \mathbf{1}_{[0;1]}(x)\,.$$

Найдем моменты ξ и η:

$$E\xi = \int\limits_{-\infty}^{+\infty} x f_\xi(x)\,dx = \int\limits_0^1 x\left(\frac{3}{2} - x\right)\,dx = \left(\frac{3x^2}{4} - \frac{x^3}{3}\right)\Big|_0^1 = \frac{5}{12}\,,$$

$$E\eta = \int\limits_{-\infty}^{+\infty} x f_\eta(x)\,dx = \int\limits_0^1 x\left(\frac{1}{2} + x\right)\,dx = \left(\frac{x^2}{4} + \frac{x^3}{3}\right)\Big|_0^1 = \frac{7}{12}\,,$$

$$E\xi^2 = \int_{-\infty}^{+\infty} x^2 f_\xi(x)\,dx = \int_0^1 x^2\left(\frac{3}{2} - x\right)\,dx = \left(\frac{x^3}{2} - \frac{x^4}{4}\right)\Big|_0^1 = \frac{1}{4}\,,$$

$$E\eta^2 = \int_{-\infty}^{+\infty} x^2 f_\eta(x)\,dx = \int_0^1 x^2\left(\frac{1}{2} + x\right)\,dx = \left(\frac{x^3}{6} + \frac{x^4}{4}\right)\Big|_0^1 = \frac{5}{12}\,,$$

$$D\xi = E\xi^2 - (E\xi)^2 = \frac{1}{4} - \frac{25}{144} = \frac{11}{144}\,,$$

$$D\eta = E\eta^2 - (E\eta)^2 = \frac{5}{12} - \frac{49}{144} = \frac{11}{144}\,.$$

Для нахождения ковариационного момента ξ и η воспользуемся формулой (C-V). Для этого найдем сначала $E(\xi\eta)$. По формуле (7.12) получим

$$E(\xi\eta) = \iint\limits_{\mathbb{R}^2} xy f_{\xi\eta}(x,y)\,dx\,dy = \int_0^1 dx \int_0^1 xy(1 - x + y)\,dy =$$

$$= \int_0^1 \left(x(1-x)\frac{1}{2} + \frac{x}{3}\right)\,dx = \int_0^1 \left(\frac{5x}{6} - \frac{x^2}{2}\right)\,dx = \frac{1}{4}\,.$$

В результате

$$\mathrm{Cov}\,(\xi,\eta) = E(\xi\eta) - E\xi\cdot E\eta = \frac{1}{4} - \frac{35}{144} = \frac{1}{144}\,.$$

Найдем коэффициент корреляции:

$$\rho_{\xi\eta} = \frac{\mathrm{Cov}\,(\xi,\eta)}{\sigma_\xi\sigma_\eta} = \frac{1}{11}\,.$$

Уравнение прямой линейной среднеквадратической регрессии η на ξ имеет вид

$$y = E\eta + \rho_{\xi\eta}\frac{\sigma_\eta}{\sigma_\xi}\,(x - E\xi) = \frac{7}{12} + \frac{1}{11}\left(x - \frac{5}{12}\right)\,.$$

Уравнение прямой линейной среднеквадратической регрессии ξ на η имеет вид

$$x = E\xi + \rho_{\xi\eta}\frac{\sigma_\xi}{\sigma_\eta}(y - E\eta) = \frac{5}{12} + \frac{1}{11}\left(y - \frac{7}{12}\right).$$

■

7.5. Задачи для самостоятельного решения

1. Найти математическое ожидание $E\xi$, дисперсию $D\xi$ и стандартное отклонение σ_ξ случайной величины ξ из задачи 2 раздела 5.6. Найти $P(|\xi - E\xi| < \sigma_\xi)$.
2. Найти $E\xi$, $D\xi$ и σ_ξ, где ξ — число попаданий в мишень стрелка из задачи 3 раздела 5.6. Выполняется ли для ξ правило двух σ?
3. В условиях задачи 4 раздела 5.6 найти математическое ожидание и дисперсию номера посетителя магазина, купившего последний экземпляр товара.
4. Найти $E\xi$, $E\eta$, $D\xi$ и $D\eta$ для случайных величин, определеннных в задаче 5 раздела 5.6. Найти $P(|\xi - E\xi| < \sigma_\xi)$ и $P(|\eta - E\eta| < \sigma_\eta)$.
5. Распределение вероятностей случайной величины ξ задано функцией распределения

$$F_\xi(x) = \begin{cases} 0\,, & x < 1\,, \\ a(x-1)^3\,, & 1 \le x < 3\,, \\ 1\,, & x \ge 3\,. \end{cases}$$

   Найти значение параметра a, $E\xi$, $D\xi$, $P(|\xi - E\xi| < 2\sigma_\xi)$.
6. Распределение вероятностей случайной величины ξ задано плотностью

$$f_\xi(x) = ax(2-x)^2 \cdot \mathbf{1}_{[0;2]}(x)\,.$$

   Найти значение параметра a, $E\xi$, $D\xi$, $P(|\xi - E\xi| < 2\sigma_\xi)$.
7. Найти ковариационный момент и коэффициент корреляции случайных величин ζ и ν из задачи 6 раздела 6.6. Построить прямые линейной среднеквадратической регрессии ζ на ν и ν на ζ.

8. Случайный вектор $\vec{\xi} = (\xi, \eta)^T$ равномерно распределен в области

$$S = \{(x, y) \,|\, 0 \le x \le 1\,,\ 0 \le y \le 2x - x^2\}\,.$$

Найти ковариационный момент и коэффициент корреляции ξ и η, построить прямые линейной среднеквадратической регрессии ξ на η и η на ξ.

9. Плотность совместного распределения случайных величин ξ и η имеет вид $f_{\xi\eta}(x, y) = a(1 - x) \cdot \mathbf{1}_S(x, y)$, где

$$S = \{(x, y) \,|\, 0 \le x \le 1\,,\ 0 \le y \le 1 - x\}\,.$$

Найти константу a, ковариационный момент и коэффициент корреляции ξ и η, построить прямые линейной среднеквадратической регрессии ξ на η и η на ξ.

Глава 8. Условное математическое ожидание

В этой главе мы рассмотрим еще одну задачу прогноза. Для ее постановки и решения нам понадобится понятие условного распределения вероятностей одной случайной величины относительно другой.

Определение 8.0.1. Пусть ξ и η — случайные величины, определенные на вероятностном пространстве $(\Omega, \mathcal{F}, P)$. ***Условным распределением вероятностей*** случайной величины ξ при условии $\eta = y$ называется вероятностная мера $P_{\xi|\eta=y}$, определенная на борелевской сигма-алгебре $\mathcal{B}(\mathbb{R})$ равенством

$$P_{\xi|\eta=y}(B) = P\left(\xi \in B \,|\eta = y\right)\,, \quad B \in \mathcal{B}(\mathbb{R})\,.$$

Мы рассмотрим способы описания $P_{\xi|\eta=y}$ в двух наиболее важных частных случаях — для дискретных случайных величин и для абсолютно непрерывных случайных величин.

8.1. Условное распределение вероятностей и условное математическое ожидание дискретных случайных величин

Рассмотрим случайный вектор $\vec{\xi} = (\xi, \eta)^T$, где ξ и η — дискретные случайные величины, определенные на вероятностном пространстве

$(\Omega, \mathcal{F}, P)$, соответствующем некоторому случайному эксперименту. Пусть их совместное распределение задано набором вероятностей

$$p_{ij} = P(\xi = x_i^1; \eta = x_j^2)\,.$$

По формуле (6.3) вероятности маргинальных распределений ξ и η имеют вид

$$p_i^1 = P(\xi = x_i^1) = \sum_j p_{ij}\,, \quad p_j^2 = P(\eta = x_j^2) = \sum_i p_{ij}\,.$$

По определению условной вероятности относительно события с ненулевой вероятностью (Определение 2.3.1) имеем:

$$P(\xi = x_i^1 | \eta = x_j^2) = \frac{p_{ij}}{p_j^2} = \frac{p_{ij}}{\sum_i p_{ij}}\,. \tag{8.1}$$

Эти вероятности полностью определяют $P_{\xi|\eta=x_j^2}$ — условное распределение вероятностей случайной величины ξ (по значениям x_i^1) при условии, что $\eta = x_j^2$. Аналогично, вероятности

$$P(\eta = x_j^2 | \xi = x_i^1) = \frac{p_{ij}}{p_i^1} = \frac{p_{ij}}{\sum_j p_{ij}}$$

задают $P_{\eta|\xi=x_i^1}$ — условное распределение вероятностей случайной величины η (по значениям x_j^2) при условии, что $\xi = x_i^1$.

Пользуясь условными распределениями вероятностей, мы можем вычислить условные математические ожидания:

$$E(\xi|\eta = x_j^2) = \sum_i x_i^1 \frac{p_{ij}}{p_j^2}\,, \tag{8.2}$$

$$E(\eta|\xi = x_i^1) = \sum_j x_j^2 \frac{p_{ij}}{p_i^1}\,. \tag{8.3}$$

Определение 8.1.1. Случайная величина, определенная равенством

$$E(\xi|\eta)(\omega) = \sum_j E(\xi|\eta = x_j^2)\mathbf{1}_{\{\eta = x_j^2\}}(\omega)\,, \tag{8.4}$$

называется ***условным математическим ожиданием*** случайной величины ξ относительно случайной величины η.

Иначе говоря, $E(\xi|\eta)$ — случайная величина, принимающая значение $E(\xi|\eta = x_j^2)$, если в результате случайного эксперимента η приняла значение x_j^2. Таким образом, условное математическое ожидание $E(\xi|\eta)$ фактически является функцией от η.

Аналогично определяется условное математическое ожидание η относительно ξ:

$$E(\eta|\xi) = \sum_i E(\eta|\xi = x_i^1)\mathbf{1}_{\{\xi=x_i^1\}}(\omega)\,.$$

Предложение 8.1.1. Условное математическое ожидание одной дискретной случайной величины относительно другой обладает следующими свойствами:
(CE-I) $E(\xi|\eta) \in \sigma(\eta)|\mathcal{B}(\mathbb{R})$;
(CE-II) $E\big(E(\xi|\eta)\big) = E\xi$.

Доказательство. Пусть $B \in \mathcal{B}(\mathbb{R})$. Справедливо равенство

$$E(\xi|\eta)^{-1}(B) = \{\omega \in \Omega \,|\, E(\xi|\eta)(\omega) \in B\} = \bigcup_{E(\xi|\eta=x_j^2)\in B} \{\eta = x_j^2\}\,.$$

Поскольку каждое из событий $\{\eta = x_j^2\}$ принадлежит сигма-алгебре $\sigma(\eta)$, отсюда следует $E(\xi|\eta)^{-1}(B) \in \sigma(\eta)$. Таким образом, $E(\xi|\eta)$ — $\sigma(\eta)|\mathcal{B}(\mathbb{R})$-измеримая случайная величина. Свойство (CE-I) доказано. Далее,

$$E\big(E(\xi|\eta)\big) = \sum_j E(\xi|\eta = x_j^2)P(\eta = x_j^2) = \sum_j \sum_i x_i^1 \frac{p_{ij}}{p_j^2} p_j^2 = \\ = \sum_i x_i^1 \sum_j p_{ij} = \sum_i x_i^1 p_i^1 = E\xi\,,$$

что доказывает свойство (CE-II). ■

R Свойство (CE-II) называют формулой ***полного математического ожидания***. Формула полной вероятности является ее частным случаем. Действительно, пусть $\{B_j\} \subseteq \mathcal{F}$ — полная группа событий, $A \in \mathcal{F}$. Определим случайные величины

$$\xi(\omega) = \mathbf{1}_A(\omega)\,, \quad \eta_j(\omega) = \mathbf{1}_{B_j}(\omega)\,.$$

Тогда $E\xi = P(A)$, при этом $E(\xi|\eta = j) = P(A|B_j)$, следовательно,

$$E(\xi|\eta)(\omega) = \sum_j P(A|B_j)\mathbf{1}_{B_j}(\omega)\,.$$

В результате получаем

$$E\big(E(\xi|\eta)\big) = \sum_j P(A|B_j)P(B_j)\,.$$

Свойства (CE-I) и (CE-II) берутся в качестве характеристических для общего определения условного математического ожидания.

Определение 8.1.2. Пусть ξ и η — случайные величины, определенные на вероятностном пространстве $(\Omega, \mathcal{F}, P)$. Условным математическим ожиданием случайной величины ξ относительно η называется $\sigma(\eta)|\mathcal{B}(\mathbb{R})$-измеримая случайная величина $E(\xi|\eta)$, такая, что $E\left(E(\xi|\eta)\right) = E\xi$.

Оказывается, случайная величина с такими свойствами существует и единственна (формула (8.4) дает ее явное представление в случае дискретных ξ и η). Однако строгое обоснование этого факта в общем случае требует нетривиального математического аппарата из области теории меры и измеримых функций. Поэтому мы ограничимся лишь нестрогим выводом формул для вычисления плотности условного распределения вероятностей и условного математического ожидания для важного и часто встречающегося в приложениях частного случая абсолютно непрерывных случайных величин.

8.2. Условное математическое ожидание абсолютно непрерывных случайных величин

Пусть $\vec{\xi} = (\xi, \eta)^T$ — случайный вектор, где ξ и η — абсолютно непрерывные случайные величины, определенные на вероятностном пространстве $(\Omega, \mathcal{F}, P)$, $f_{\xi\eta}(x, y)$ — плотность их совместного распределения вероятностей. Как было показано в главе 6, она полностью определяет маргинальные распределения вероятностей ξ и

η. Их плотности распределения находятся по формуле (6.6). Нам нужно описать вероятностную меру $P_{\xi|\eta=y}$ — распределение вероятностей ξ при условии, что $\eta = y$. Сделаем это, построив плотность условного распределения вероятностей, которую будем обозначать $f_{\xi|\eta}(x|\eta = y)$. По аналогии с тем, как это делается при построении плотности распределения вероятностей в безусловном случае, найдем сначала функцию распределения ξ при условии $\eta = y$:

$$F_{\xi|\eta}(x|\eta = y) := P(\xi < x|\eta = y)\,.$$

Проблема здесь в том, что условная вероятность в правой части этого равенства не определена в смысле определения 2.3.1. Это связано с тем, что вероятность условия $\{\eta = y\}$ здесь равна нулю. Действительно, для абсолютно непрерывных случайных величин вероятность любого из их возможных значений равна нулю. В качестве выхода из положения используем предельный переход, а именно положим по определению

$$P(\xi < x|\eta = y) := \lim_{\Delta y\to 0} P(\xi < x|\eta \in [y; y + \Delta y])\,,$$

для тех $y \in \mathbb{R}$, для которых события вида $\{\eta \in [y; y + \Delta y]\}$ имеют ненулевую вероятность при малых Δy.

$$P(\xi < x|\eta \in [y; y + \Delta y]) = \frac{P(\xi < x; \eta \in [y; y + \Delta y])}{P(\{\eta \in [y; y + \Delta y\})} =$$
$$= \frac{\displaystyle\int_{-\infty}^{x} du \int_{y}^{y+\Delta y} f_{\xi\eta}(u, v)\, dv}{\displaystyle\int_{y}^{y+\Delta y} f_{\eta}(v)\, dv} = \frac{\displaystyle\int_{y}^{y+\Delta y} dv \int_{-\infty}^{x} f_{\xi\eta}(u, v)\, du}{\displaystyle\int_{y}^{y+\Delta y} dv \int_{-\infty}^{+\infty} f_{\xi\eta}(u, v)\, du} =$$
$$= \frac{\displaystyle\int_{-\infty}^{x} f_{\xi\eta}(u, y + \theta_1\Delta y)\, du \cdot \Delta y}{\displaystyle\int_{-\infty}^{+\infty} f_{\xi\eta}(u, y + \theta_2\Delta y)\, du \cdot \Delta y}\,,$$

где θ_1 и θ_2 — некоторые константы из интервала $(0; 1)$, здесь мы применили интегральную теорему о среднем к интегралам по переменной v в числителе и знаменателе дроби. Сокращая множитель Δy

и переходя к пределу в полученном равенстве при $\Delta y \to 0$, получаем

$$F_{\xi|\eta}(x|\eta = y) := P(\xi < x|\eta = y) = \frac{\displaystyle\int_{-\infty}^{x} f_{\xi\eta}(u, y)\, du}{\displaystyle\int_{-\infty}^{+\infty} f_{\xi\eta}(u, y)\, du}.$$

Дифференцируя полученное равенство по x, получаем формулу для плотности условного распределения вероятностей:

$$\boxed{f_{\xi|\eta}(x|\eta = y) = \frac{f_{\xi\eta}(x, y)}{\displaystyle\int_{-\infty}^{+\infty} f_{\xi\eta}(x, y)\, dx}. \qquad (8.5)}$$

Пользуясь полученной формулой, можем получить формулу для условного математического ожидания ξ при условии $\eta = y$:

$$\boxed{E(\xi|\eta = y) = \int_{-\infty}^{+\infty} x f_{\xi|\eta}(x|\eta = y)\, dx = \frac{\displaystyle\int_{-\infty}^{+\infty} x f_{\xi\eta}(x, y)\, dx}{\displaystyle\int_{-\infty}^{+\infty} f_{\xi\eta}(x, y)\, dx}. \qquad (8.6)}$$

Как следствие, для абсолютно непрерывных случайных величин мы получаем явное представление условного математического ожидания $E(\xi|\eta)$ как функции η:

$$\boxed{E(\xi|\eta) = \frac{\displaystyle\int_{-\infty}^{+\infty} x f_{\xi\eta}(x, \eta)\, dx}{\displaystyle\int_{-\infty}^{+\infty} f_{\xi\eta}(x, \eta)\, dx}. \qquad (8.7)}$$

R Формула (8.5) для плотности условного распределения вероятностей абсолютно непрерывных случайных величин устроена аналогично формуле (8.1) для вероятностей условного распределения дискретных случайных величин. Роль p_{ij} играет $f_{\xi\eta}(x, y)$, а роль суммирования по i в знаменателе дроби играет интегрирование по x. Та же аналогия наблюдается в формулах (8.2), (8.6) и в формулах (8.4), (8.7).

8.3. Среднеквадратическая регрессия

В предыдущей главе была рассмотрена простейшая задача прогнозирования и получена формула линейной среднеквадратической регрессии одной случайной величины на другую. С помощью условного математического ожидания решается ***задача о (нелинейной) среднеквадратической регрессии***. Сформулируем постановку этой задачи.

Пусть $\vec{\xi} = (\xi, \eta)^T$ — случайный вектор, где ξ и η — случайные величины, определенные на вероятностном пространстве $(\Omega, \mathcal{F}, P)$, которое отвечает некоторому случайному эксперименту. Предположим, что в результате проведения этого случайного эксперимента η приняла значение y. Необходимо найти наилучший в среднеквадратичном прогноз значения ξ по наблюдаемому значению η, т. е. такое $g(y)$, что

$$E\big((\xi - g(y))^2 \big| \eta = y\big) = \min_{a \in \mathbb{R}} E\big((\xi - a)^2 \big| \eta = y\big) .$$

Решение поставленной задачи получается с помощью свойства (D-III) (стр. 115) математического ожидания и дисперсии. Очевидно, это свойство останется верным при переходе от математического ожидания и дисперсии, вычисляемых относительно исходной вероятностной меры P, к математическому ожиданию и дисперсии, вычисляемых относительно условного распределения вероятностей $P_{\xi|\eta=y}$. В результате

$$\min_{a \in \mathbb{R}} E\big((\xi - a)^2 \big| \eta = y\big) = E\big([\xi - E(\xi|\eta = y)]^2 \big| \eta = y\big) ,$$

т. е. решением поставленной задачи является условное математическое ожидание:

$$g(y) = E(\xi|\eta = y) . \tag{8.8}$$

Функция $x = g(y)$, определенная формулой (8.8), называется ***кривой среднеквадратической регрессии случайной величины $\boldsymbol{\xi}$ на случайную величину $\boldsymbol{\eta}$***. Аналогично функция $y = h(x)$, где

$$h(x) = E(\eta|\xi = x)\,,$$

называется ***кривой среднеквадратической регрессии случайной величины η на случайную величину ξ***.

Задача 8.3.1. Совместное распределение вероятностей случайных величин ξ и η задано плотностью (см. задачу 7.4.1)

$$f_{\xi\eta}(x,y) = (1 - x + y)\cdot \mathbf{1}_{[0;1]\times[0;1]}(x,y)\,.$$

Построить кривые среднеквадратической регрессии случайных величин ξ и η друг на друга.

Решение. Используя найденные при решении задачи 7.4.1 плотности маргинальных распределений вероятностей

$$f_\xi(x) = \int\limits_{-\infty}^{+\infty} f_{\xi\eta}(x,y)\,dy = \left(\frac{3}{2} - x\right)\cdot \mathbf{1}_{[0;1]}(x)$$

и

$$f_\eta(y) = \int\limits_{-\infty}^{+\infty} f_{\xi\eta}(x,y)\,dx = \left(\frac{1}{2} + y\right)\cdot \mathbf{1}_{[0;1]}(y)\,,$$

по формуле (8.6) находим кривую среднеквадратической регрессии ξ на η:

$$x = g(y) = E(\xi|\eta = y) = \frac{\int\limits_{-\infty}^{+\infty} x f_{\xi\eta}(x,y)\,dx}{\int\limits_{-\infty}^{+\infty} f_{\xi\eta}(x,y)\,dx} =$$

$$= \frac{\int\limits_0^1 x(1 - x + y)\,dx}{\dfrac{1}{2} + y} = \frac{3y+1}{3(2y+1)}\,, \quad y \in [0;1]\,.$$

Аналогично находим кривую среднеквадратической регрессии η на ξ:

$$y = h(x) = E(\eta|\xi = x) = \frac{\int\limits_{-\infty}^{+\infty} y f_{\xi\eta}(x, y)\, dy}{\int\limits_{-\infty}^{+\infty} f_{\xi\eta}(x, y)\, dy} =$$

$$= \frac{\int\limits_{0}^{1} y(1 - x + y)\, dy}{\frac{3}{2} - x} = \frac{5 - 3x}{3(3 - 2x)}, \quad x \in [0; 1]$$

Таким образом, формулы

$$E(\xi|\eta) = \frac{3\eta + 1}{3(2\eta + 1)}; \quad E(\eta|\xi) = \frac{5 - 3\xi}{3(3 - 2\xi)}$$

дают наилучший в среднеквадратическом прогноз значения одной компоненты случайного вектора $\vec{\xi}$ по наблюдаемому значению другой. ■

8.4. Задачи для самостоятельного решения

1. Для случайных величин ζ и ν из задачи 6 раздела 6.6 построить кривые среднеквадратической регрессии ζ на ν и ν на ζ. Сравнить полученный результат с результатом задачи 7 раздела 7.5.
2. Для случайных величин ξ и η задачи 8 раздела 7.5 построить кривые среднеквадратичной регрессии ξ на η и η на ξ. Сравнить полученный результат с результатом задачи 8 раздела 7.5.
3. Для случайных величин ξ и η задачи 9 раздела 7.5 построить кривые среднеквадратичной регрессии ξ на η и η на ξ. Сравнить полученный результат с результатом задачи 9 раздела 7.5.

Глава 9. Характеристические функции

Характеристические функции представляют собой удобный математический аппарат для изучения распределений вероятностей. В частности, с их помощью удобно исследовать распределения вероятностей сумм независимых случайных величин и пределов последовательностей случайных величин.

9.1. Определение характеристической функции

Определение 9.1.1. ***Характеристической функцией*** случайной величины ξ, определенной на некотором вероятностном пространстве $(\Omega, \mathcal{F}, P)$ (или распределения вероятностей P_ξ), называется комплекснозначная функция действительной переменной, определенная равенством

$$\varphi_\xi(t) = Ee^{i\xi t}\,.$$

Рассмотрим два примера, иллюстрирующих связь между характеристической функцией и распределением вероятностей случайной величины.

■ Пример 9.1.1. Пусть ξ — дискретная случайная величина, принимающая целые неотрицательные значения. Ее распределение вероятностей полностью определяется вероятностями

$$p_k = P(\xi = k)\,, \quad k = 0, 1, 2, \ldots\,.$$

Для характеристической функции ξ получаем представление

$$\varphi_\xi(t) = Ee^{i\xi t} = \sum_{k=0}^{\infty} e^{ikt} p_k\,.$$

Рассмотрим действительную часть этой функции:

$$\begin{aligned}\operatorname{Re}\varphi_\xi(t) &= \frac{1}{2}\left(\varphi(t) + \overline{\varphi(t)}\right) = \\ &= \frac{1}{2}\left(\sum_{k=0}^{\infty} e^{ikt} p_k + \sum_{k=0}^{\infty} e^{-ikt} p_k\right) = \sum_{k=-\infty}^{\infty} e^{ikt} c_k\,,\end{aligned}$$

где

$$c_k = \begin{cases} p_0\,, & k = 0\,,\\ \dfrac{p_k}{2}\,, & k > 0\,,\\ \dfrac{p_{-k}}{2}\,, & k < 0\,.\end{cases}$$

Этот ряд является тригонометрическим рядом Фурье своей суммы, поэтому, по известной формуле коэффициентов ТРФ в комплексной форме, получаем:

$$p_k = 2c_k = \frac{2}{\pi}\int_{-\pi}^{\pi} \operatorname{Re}\varphi_\xi(t) e^{-ikt} dt\,, \quad p_0 = c_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} \operatorname{Re}\varphi_\xi(t)\, dt\,.$$

Таким образом, характеристическая функция полностью определяет ряд распределения данной дискретной случайной величины. ■

■ Пример 9.1.2. Пусть ξ — абсолютно непрерывная случайная величина, $f_\xi(x)$ — ее плотность распределения вероятностей. Тогда $\varphi_\xi(t)$ и $f_\xi(x)$ связаны друг с другом преобразованием Фурье:

$$\varphi_\xi(t) = Ee^{i\xi t} = \int_{-\infty}^{+\infty} e^{ixt} f_\xi(x)\, dx\,.$$

С помощью обратного преобразования Фурье получаем:

$$f_\xi(x) = \frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} e^{-ixt} \varphi_\xi(t)\, dt\,.$$

Таким образом, распределение вероятностей абсолютно непрерывной случайной величины также полностью восстанавливается по ее характеристической функции. ■

R В классическом математическом анализе прямое преобразование Фурье функции $f(t)$ определяется равенством

$$F(x) = \mathcal{F}[f](x) = \int\limits_{-\infty}^{+\infty} e^{-ixt} f(t)\, dt\,,$$

а обратное — равенством

$$f(t) = \mathcal{F}^{-1}[F](t) = \frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} e^{ixt} F(x)\, dx\,.$$

В соответствии со сложившейся традицией, в теории вероятностей для определения характеристической функции случайной величины фактически используется (с точностью до множителя $1/2\pi$) обратное преобразование Фурье.

Оказывается, связь между характеристической функцией и распределением вероятностей случайной величины имеет место в самом общем случае: ***характеристическая функция полностью определяет распределение вероятностей случайной величины***. Это следует, например, из следующей теоремы, которую мы приведем без доказательства.

Теорема 9.1.1. — Формула обращения Леви. Пусть ξ — случайная величина, распределение вероятностей которой задано функцией распределения $F_\xi(x)$, $\varphi_\xi(t)$ — ее характеристическая функция. Тогда для любых $x, y \in \mathbb{R}$

$$F_\xi(y) - F_\xi(x) = \frac{1}{2\pi} v.p. \int\limits_{-\infty}^{+\infty} \frac{e^{-ixt} - e^{-iyt}}{it} \varphi_\xi(t)\, dt\,. \tag{9.1}$$

Формула (9.1) определяет вероятности событий вида $\{\xi \in [x; y)\}$, т. е. вероятностную меру P_ξ (распределение вероятностей случайной величины ξ) на промежутках вида $[x; y)$. Оказывается, это однозначно определяет P_ξ на всей порожденной этой совокупностью множеств сигма-алгебре $\mathcal{B}(\mathbb{R})$ (факт, который мы примем без доказательства).

Взаимно однозначное соответствие между распределениями вероятностей случайных величин и характеристическими функциями лежит в основе их использования для изучения распределений вероятностей.

9.2. Свойства характеристических функций

Характеристические функции обладают следующими свойствами.

(φ-I) $\varphi_\xi(t)$ определена при всех $t \in \mathbb{R}$, при этом
$\varphi_\xi(0) = 1$ и $\forall\, t \in \mathbb{R}\ \ |\varphi_\xi(t)| \le 1$;

(φ-II) $\forall\, a, b \in \mathbb{R}\ \ \varphi_{a\xi+b} = e^{ibt}\varphi_\xi(at)$;

(φ-III) $E\xi = -i\varphi'_\xi(0)$, если математическое ожидание существует;

(φ-IV) $D\xi = -\varphi''_\xi(0) + \big(\varphi'_\xi(0)\big)^2$, если дисперсия существует;

(φ-V) Если случайные величины $\xi_1, \xi_2, \dots, \xi_n$ независимы, то
$$\varphi_{\xi_1+\dots+\xi_n}(t) = \prod_{k=1}^{n} \varphi_{\xi_k}(t)\,.$$

(φ-VI) $\varphi_\xi(t)$ положительно-определенная функция, т. е.
$$\forall\, n \in \mathbb{N}\ \ t_1, \dots, t_n \in \mathbb{R}, z_1, \dots, z_n \in \mathbb{C} \sum_{i,j=1}^{n} \varphi(t_i - t_j) z_i \overline{z}_j \ge 0\,.$$

Доказательство. Из оценки $|e^{i\xi t}| \le 1$ следует существование $Ee^{i\xi t}$ при всех $t \in \mathbb{R}$, кроме того, в силу свойства (E-IV), $|Ee^{i\xi t}| \le E1 = 1$. Непосредственной подстановкой $t = 0$ получаем $\varphi_\xi(0) = Ee^0 = 1$.

Доказательство свойства (φ-II) получается с помощью свойства линейности математического ожидания:

$$\varphi_{a\xi+b} = Ee^{(a\xi+b)t} = e^{ibt}Ee^{i\xi at} = e^{ibt}\varphi_\xi(at)\,.$$

Для доказательства свойств (φ-III) и (φ-IV) продифференцируем характеристическую функцию:

$$\varphi'_\xi(t) = E\left[i\xi e^{i\xi t}\right] = iE\left[\xi e^{i\xi t}\right]\,, \quad \varphi''_\xi(t) = -E\left[\xi^2 e^{i\xi t}\right]\,.$$

Подставляя в полученные равенства $t = 0$, получаем

$$\varphi'_\xi(0) = iE\xi\,, \quad \varphi''_\xi(0) = -E\xi^2\,.$$

Из первого равенства сразу получается свойство (φ-III), а по формуле (D-II) (стр. 115) из полученных равенств получается свойство (φ-IV).

Если $\xi_1, \xi_2, \ldots, \xi_n$ — независимые случайные величины, функции $e^{i\xi_1 t}, e^{i\xi_2 t}, \ldots, e^{i\xi_n t}$ от этих величин также являются независимыми случайными величинами, поэтому, по теореме 7.1.10 о математическом ожидании произведения независимых случайных величин, получаем

$$\varphi_{\xi_1+\cdots+\xi_n}(t) = Ee^{i(\xi_1+\cdots+\xi_n)t} = E\prod_{k=1}^n e^{i\xi_k t} = \prod_{k=1}^n Ee^{i\xi_k t} = \prod_{k=1}^n \varphi_{\xi_k}(t)\,.$$

Для любых $n \in \mathbb{N}, t_1, \ldots, t_n \in \mathbb{R}, z_1, \ldots, z_n \in \mathbb{C}$ получаем

$$\sum_{i,j=1}^n \varphi(t_i - t_j) z_i \overline{z}_j = \sum_{i,j=1}^n Ee^{i\xi(t_i-t_j)} z_i \overline{z}_j = \sum_{i,j=1}^n E\left[e^{i\xi t_i} z_i \cdot \overline{e^{i\xi t_j} z_j}\right] =$$

$$= E\sum_{i,j=1}^n \left[e^{i\xi t_i} z_i \cdot \overline{e^{i\xi t_j} z_j}\right] = E\left[\sum_{i=1}^n e^{i\xi t_i} z_i \cdot \overline{\sum_{j=1}^n e^{i\xi t_j} z_j}\right] =$$

$$= E\left|\sum_{i=1}^n e^{i\xi t_i} z_i\right|^2 \geq 0\,. \qquad \blacksquare$$

Приведем без доказательства теорему, которая дает необходимое и достаточное условие, которому должна удовлетворять функция, чтобы быть характеристической для некоторого распределения вероятностей.

Теорема 9.2.1. — Бохнер–Хинчин. Пусть $\varphi(t)$ — непрерывная функция, такая, что $\varphi(0) = 1$. Для того чтобы она была характеристической для некоторого распределения вероятностей P_ξ, необходимо и достаточно, чтобы она была ***положительно определенной***, т. е. для любого $n \in \mathbb{N}$ и для любых $t_1, \dots, t_n \in \mathbb{R}$ и $z_1, \dots, z_n \in \mathbb{C}$ выполнялось условие

$$\sum_{i,j=1}^{n} \varphi(t_i - t_j) z_i \overline{z}_j \geq 0\,.$$

■ Пример 9.2.1. Пусть $c \in \mathbb{R}$, тогда $\varphi_c(t) = Ee^{ict} = e^{ict}$. ■

■ Пример 9.2.2. Пусть $\xi \sim N(a, \sigma^2)$, тогда $\varphi_\xi(t) = e^{iat - \frac{\sigma^2 t^2}{2}}$.

Действительно, если $\xi \sim N(0, 1)$, то

$$\varphi_\xi(t) = Ee^{i\xi t} = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} e^{ixt - \frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} e^{-\frac{1}{2}\left((x-it)^2 + t^2\right)} dx =$$

$$= e^{-\frac{t^2}{2}} \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty - it}^{+\infty - it} e^{-\frac{z^2}{2}} dz = e^{-\frac{t^2}{2}} \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dz = e^{-\frac{t^2}{2}}\,.$$

Пусть теперь $\xi \sim N(a, \sigma^2)$. Пользуясь предложением 5.5.3, представим ξ в виде $\xi = \sigma\eta + a$, где $\eta \sim N(0, 1)$. Тогда, по свойству (φ-II), получаем

$$\varphi_\xi(t) = \varphi_{\sigma\eta + a}(t) = e^{iat} \varphi_\eta(\sigma t) = e^{iat - \frac{\sigma^2 t^2}{2}}\,.$$ ■

Пользуясь характеристическими функциями, можно легко доказывать различные свойства случайных величин.

■ Пример 9.2.3. Докажем следующее свойство гауссовских случайных величин:

Пусть $\xi_1 \sim N(a_1, \sigma_1)$ и $\xi_2 \sim N(a_2, \sigma_2)$ независимы, тогда

$$\forall\, b_1, b_2 \in \mathbb{R} \quad b_1\xi_1 + b_2\xi_2 \sim N\left(b_1 a_1 + b_2 a_2, b_1^2 \sigma_1^2 + b_2^2 \sigma_2^2\right)\,.$$

Для того чтобы доказать это свойство, построим характеристическую функцию случайной величины $b_1\xi_1 + b_2\xi_2$. Пользуясь свойствами (φ-V) и (φ-II) и результатом примера 9.2.2, получим:

$$\varphi_{b_1\xi_1 + b_2\xi_2}(t) = \varphi_{b_1\xi_1}(t) \varphi_{b_2\xi_2}(t) = \varphi_{\xi_1}(b_1 t) \varphi_{\xi_2}(b_2 t) =$$

$$= e^{ia_1t-\frac{\sigma_1^2 b_1^2 t^2}{2}} e^{ia_2t-\frac{\sigma_2^2 b_2^2 t^2}{2}} = e^{i(a_1+a_2)t-\frac{(b_1^2\sigma_1^2+b_2^2\sigma_2^2)t^2}{2}}.$$

В силу взаимной однозначности соответствия между распределениями вероятностей и характеристическими функциями, из полученного равенства следует, что $b_1\xi_1+b_2\xi_2$ — гауссовская случайная величина с математическим ожиданием a_1+a_2 и дисперсией $b_1^2\sigma_1^2+b_2^2\sigma_2^2$, что и требовалось доказать. ■

С помощью характеристических функций можно строить новые распределения вероятностей, т. е. получать новые математические модели случайных явлений.

■ Пример 9.2.4. — Гамма-распределение. В примере 5.5.2 была рассмотрена одна из математических моделей теории надежности. Было показано, что показательное распределение вероятностей описывает время ожидания при отсутствии последействия. Построим модель надежности, учитывающую такие явления, как износ, старение и т. п.

Пусть ξ — время исправной работы некоторого агрегата, работоспособность которого зависит от узла, который может выйти из строя. При этом есть возможность мгновенной замены этого узла на такой же новый и имеется k запасных узлов. Пусть ξ_j, где $j=1,2,\ldots,k$ — время исправной работы j-го узла. Будем предполагать, что $\xi_j \sim \mathrm{Exp}(\lambda)$. Тогда $\theta := E\xi_j = 1/\lambda$ — среднее время работы j-го узла. Характеристическая функция ξ_j имеет вид

$$\varphi_{\xi_j}(t) = Ee^{i\xi_j t} = \lambda \int\limits_0^{+\infty} e^{ixt}e^{-\lambda x}dx = -\frac{\lambda e^{-(\lambda-it)x}}{\lambda-it}\Bigg|_0^{+\infty} = \frac{\lambda}{\lambda-it}.$$

Поскольку $\xi=\sum\limits_{j=1}^{k}\xi_j$, а случайные величины ξ_j независимы, по свойству (φ-V), характеристическая функция случайной величины ξ имеет вид:

$$\varphi_\xi(t) = \frac{\lambda^k}{(\lambda-it)^k}.$$

В результате мы можем найти плотность распределения вероятностей ξ (см. пример 9.1.2):

$$f_\xi(x) = \frac{1}{2\pi}\int\limits_{-\infty}^{+\infty} e^{-ixt}\varphi_\xi(t)\,dt = \frac{\lambda^k}{2\pi}\int\limits_{-\infty}^{+\infty}\frac{e^{-ixt}}{(\lambda - it)^k}\,dt =$$

$$= \frac{\lambda^k}{2\pi i^k}\int\limits_{-\infty}^{+\infty}\frac{e^{-ixt}}{(-t - i\lambda)^k}\,dt = \frac{\lambda^k}{2\pi i^k}\int\limits_{-\infty}^{+\infty}\frac{e^{ixz}}{(z - i\lambda)^k}\,dz\,.$$

Последний интеграл получен в результате замены $-t = z$. Его можно вычислить с помощью вычетов:

$$\int\limits_{-\infty}^{+\infty}\frac{e^{ixz}}{(z - i\lambda)^k}\,dz = 2\pi i \operatorname*{res}_{z=i\lambda}\frac{e^{ixz}}{(z - i\lambda)^k} =$$

$$= 2\pi i\,\frac{1}{(k-1)!}\lim_{z\to i\lambda}\frac{d^{k-1}}{dz^{k-1}}\,e^{ixz} = \frac{2\pi i^k x^{k-1}}{(k-1)!}e^{-\lambda x}\,.$$

Здесь мы воспользовались тем, что $z = i\lambda$ — полюс порядка k функции $\dfrac{e^{ixz}}{(z - i\lambda)^k}$. Таким образом,

$$f_\xi(x) = \frac{\lambda^k x^{k-1}}{(k-1)!}e^{-\lambda x} = \frac{x^{k-1}}{\theta^k\Gamma(k)}e^{-\frac{x}{\theta}}\,,$$

где Γ — гамма-функция Эйлера (см. Приложение В).

Полученное распределение вероятностей называется ***гамма-распределение***.

Обозначение: $\boxed{\xi \sim \Gamma(k,\theta)}$

Определение 9.2.1. Случайная величина ξ ***распределена по закону гамма с параметрами $k > 0$, $\theta > 0$*** (k необязательно целое), если ее плотность распределения имеет вид

$$f_\xi(x) = \frac{x^{k-1}}{\theta^k\Gamma(k)}e^{-\frac{x}{\theta}}\cdot\mathbf{1}_{[0;+\infty)}(x)\,. \tag{9.2}$$

При этом, как было показано выше, ее характеристическая функция имеет вид

$$\varphi_\xi(t) = \frac{1}{(1 - i\theta t)^k}\,. \tag{9.3}$$

■

9.3. Задачи для самостоятельного решения

1. Пусть $\xi \sim \Gamma(k, \theta)$. Найти $E\xi$ и $D\xi$ двумя способами: пользуясь свойствами гамма-функции Эйлера (см. Приложение B) и используя характеристическую функцию.
2. Найти распределения вероятностей, которым соответствуют следующие характеристические функции:

   a) $\varphi(t) = \cos^2 t\,;$

   b) $\varphi(t) = \dfrac{1}{2}(\cos t + \cos 2t)\,;$

   c) $\varphi(t) = \dfrac{2}{3} + \dfrac{1}{3}\cos 3t\,.$
3. Найти характеристическую функцию случайной величины ξ, имеющей треугольное распределение вероятностей, заданное плотностью $f_\xi(x) = a(1 - a|x|) \cdot \mathbf{1}_{\left[-\frac{1}{a};\frac{1}{a}\right]}(x)$.
4. Найти характеристическую функцию случайной величины ξ, имеющей плотность распределения вероятностей $f_\xi(x) = 2x \cdot \mathbf{1}_{[0;1]}(x)$. Найти $E\xi$ с помощью характеристической функции.
5. Найти плотность распределения вероятностей случайной величины ξ, имеющей характеристическую функцию

   $$\varphi_\xi(t) = \frac{1}{1+t^2}\,.$$

   Найти $E\xi$ и $D\xi$ с помощью характеристической функции.
6. Пусть ξ_k — индикатор успеха в k-м испытании Бернулли с вероятностью успеха p.

   a) Найти $\varphi_{\xi_k}(t)$.

   b) Пользуясь свойством (φ-V) характеристических функций, построить $\varphi_{\eta_n}(t)$, где $\eta_n = \displaystyle\sum_{k=1}^{n} \xi_k$ — число успехов в серии из n испытаний Бернулли.

   c) Используя $\varphi_{\eta_n}(t)$, вывести формулу Бернулли для вероятности $P_n(m)$ того, что в серии из n испытаний число успехов равно m.

Глава 10. Предельные теоремы теории вероятностей

В этой главе мы рассмотрим несколько теорем, характеризующих пределы последовательностей случайных величин. Для этого сначала рассмотрим различные виды сходимостей таких последовательностей.

10.1. Сходимость последовательностей случайных величин

Пусть $\{\xi_n\}$ — последовательность случайных величин, определенных на вероятностном пространстве $(\Omega, \mathcal{F}, P)$, ξ — некоторая случайная величина, определенная на том же вероятностном пространстве. Приведем определения трех наиболее часто использующихся видов сходимости.

Определение 10.1.1. Последовательность $\{\xi_n\}$ сходится к случайной величине ξ ***с вероятностью единица***, или ***почти наверное***, если

$$P\left(\left\{\omega \,\middle|\, \xi_n(\omega) \xrightarrow[n\to\infty]{} \xi(\omega)\right\}\right) = 1\,.$$

При этом пишут $\xi_n \xrightarrow[n\to\infty]{\text{п.н.}} \xi$.

Определение 10.1.2. Последовательность $\{\xi_n\}$ сходится к случайной величине ξ ***по вероятности***, если

$$\forall\, \varepsilon > 0 \quad P\left(|\xi_n - \xi| > \varepsilon\right) \xrightarrow[n\to\infty]{} 0\,.$$

При этом пишут $\xi_n \xrightarrow[n\to\infty]{P} \xi$.

Определение 10.1.3. Последовательность $\{\xi_n\}$ сходится к случайной величине ξ ***в среднем степени*** p, если

$$E|\xi_n - \xi|^p \xrightarrow[n\to\infty]{} 0\,.$$

При этом пишут $\xi_n \xrightarrow[n\to\infty]{L_p} \xi$. При $p = 2$ определенную таким образом сходимость называют ***сходимостью в среднеквадратичном***.

Связи между сходимостями последовательности случайных величин в среднеквадратичном, почти наверное и по вероятности показаны на рис. 10.1.

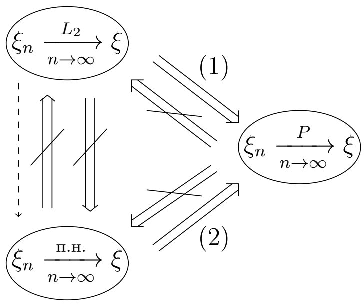


Рис. 10.1. Связи между сходимостями в среднеквадратичном, почти наверное и по вероятности

Доказательство. Импликация (1) на этой диаграмме сразу следует из неравенства Чебышева:

$$\forall\, \varepsilon > 0 \quad P\left(|\xi_n - \xi| > \varepsilon\right) \le \frac{E|\xi_n - \xi|^2}{\varepsilon^2}\,.$$

Докажем импликацию (2). Пусть $\xi_n \xrightarrow[n\to\infty]{\text{п.н.}} \xi$. По определению предела по Коши получаем $\omega \in \left\{\omega \in \Omega \mid \lim\limits_{n\to\infty} \xi_n(\omega) = \xi(\omega)\right\}$ тогда и только тогда, когда выполнено условие

$$\forall\, m \in \mathbb{N}\ \exists\, N \in \mathbb{N}\ \forall\, n \geq N\ |\xi_n(\omega) - \xi(\omega)| < \frac{1}{m}\,.$$

Отсюда следует равенство множеств (событий)

$$\Big\{\omega \in \Omega\,|\, \lim_{n\to\infty} \xi_n(\omega) = \xi(\omega)\Big\} =$$
$$= \bigcap_{m\in\mathbb{N}} \bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} \left\{\omega \in \Omega\,|\, |\xi_n(\omega) - \xi(\omega)| < \frac{1}{m}\right\} =: A\,.$$

Сходимость последовательности $\{\xi_n\}$ к ξ почти наверное означает, что вероятность события A равна 1. Очевидно,

$$A \subseteq \bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} \left\{\omega \in \Omega\,|\, |\xi_n(\omega) - \xi(\omega)| < \frac{1}{m}\right\}\,,$$

следовательно,

$$P(A) \leq P\left(\bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} \left\{\omega \in \Omega\,|\, |\xi_n(\omega) - \xi(\omega)| < \frac{1}{m}\right\}\right)\,,$$

значит,

$$P\left(\bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} \left\{\omega \in \Omega\,|\, |\xi_n(\omega) - \xi(\omega)| < \frac{1}{m}\right\}\right) = 1\,,$$

а для противоположного события получаем

$$P\left(\bigcap_{N\in\mathbb{N}} \bigcup_{n\geq N} \left\{\omega \in \Omega\,|\, |\xi_n(\omega) - \xi(\omega)| \geq \frac{1}{m}\right\}\right) = 0\,.$$

События $B_{N,m} := \bigcup_{n\geq N} \left\{\omega \in \Omega\,|\, |\xi_n(\omega) - \xi(\omega)| \geq \frac{1}{m}\right\}$ образуют убывающую (в теоретико-множественном смысле) последовательность $\{B_{N,m}\}_{N=1}^{\infty}$, поскольку $B_{N,m} \supseteq B_{N+1,m}$ для любого $N \in \mathbb{N}$. Отсюда, по свойству непрерывности вероятности (свойство (P-9)), получаем $\lim\limits_{N\to\infty} P\left(B_{N,m}\right) = P\left(\bigcap\limits_{N\in\mathbb{N}} B_{N,m}\right) = 0$. Поскольку справедливо вложение

$$\left\{\omega \in \Omega\,|\, |\xi_n(\omega) - \xi(\omega)| > \frac{1}{m}\right\} \subseteq B_{N,m}\,,$$

отсюда следует, что

$$\forall\, m \in \mathbb{N} \quad \lim_{N\to\infty} P\left(\left\{\omega \in \Omega \mid |\xi_n(\omega) - \xi(\omega)| > \frac{1}{m}\right\}\right) = 0\,.$$

Таким образом, $\xi_n \xrightarrow[n\to\infty]{P} \xi$. ■

Следующий контрпример показывает, что импликация, обратная к импликации (2) на рисунке 10.1, не верна.

▪ Пример 10.1.1. Пусть $(\Omega, \mathcal{F}, P)$ — вероятностное пространство, где $\Omega = [0;1]$, $\mathcal{F}$ — сигма-алгебра измеримых по Лебегу подмножеств $[0;1]$, P — мера Лебега на прямой. Рассмотрим случайные величины, определенные следующим образом:

$$\eta_{n,k}(\omega) = \mathbf{1}_{\left[\frac{k-1}{2^n};\frac{k}{2^n}\right)}(\omega)\,, n \in \mathbb{N}\,, \quad k = 1\ldots, 2^n\,.$$

Построим из них последовательность $\{\xi_n\}$, упорядочив следующим образом:

$$\eta_{1,1}, \eta_{1,2}, \eta_{2,1}, \eta_{2,2}, \ldots, \eta_{2,4}, \eta_{3,1}, \eta_{3,2}, \ldots, \eta_{3,8}, \ldots, \eta_{n,1}\eta_{n,2}, \ldots, \eta_{n,2^n}, \ldots\,. \tag{10.1}$$

Поскольку для любого $\varepsilon > 0$ выполнено равенство

$$P\left(\{\omega \in \Omega \mid |\eta_{n,k}(\omega)| > \varepsilon\}\right) = \frac{1}{2^n}\,,$$

последовательность ξ_n сходится по вероятности к $\xi = 0$. Однако для любого $\omega \in \Omega$ последовательность $\{\xi_n(\omega)\}$ не сходится к нулю, так как каким бы большим ни был номер n, найдутся члены этой последовательности с номерами большими, чем n, равные 1. Таким образом, последовательность не является сходящейся почти наверное.

Заметим, что $E|\eta_{n,k}|^2 = P\left(\eta_{n,k} = 1\right) = \dfrac{1}{2^n}$, поэтому последовательность $\{\xi_n\}$ сходится к $\xi = 0$ в среднеквадратичном, так что этот же пример показывает, что ***из сходимости в среднеквадратичном, вообще говоря, не следует сходимость почти наверное***. ▪

Следующий пример показывает, что импликация, обратная к импликации (1) на рисунке 10.1, не верна.

■ Пример 10.1.2. Рассмотрим последовательность $\{\xi_n\}$, заданную так же, как в примере 10.1.1 (с помощью правила (10.1)), но где

$$\eta_{n,k}(\omega) = 2^n \cdot \mathbf{1}_{\left[\frac{k-1}{2^n};\frac{k}{2^n}\right)}(\omega)\,, n \in \mathbb{N}\,, \quad k = 1\ldots, 2^n\,.$$

Так же, как в предыдущем примере, полученная последовательность $\{\xi_n\}$ сходится по вероятности к $\xi = 0$, но при этом $E|\eta_{n,k}|^2 = 2^{2n} \cdot P\left(\eta_{n,k} = 2^n\right) = 2^n$, поэтому она не сходится к $\xi = 0$ в среднеквадратичном. ■

Следующий пример показывает, что ***из сходимости почти наверное, вообще говоря, не следует сходимость в среднеквадратичном***.

■ Пример 10.1.3. Пусть $(\Omega, \mathcal{F}, P)$ то же, что в примере 10.1.1. Положим

$$\xi_n(\omega) = \sqrt{n} \cdot \mathbf{1}_{\left[0;\frac{1}{n}\right]}(\omega)\,.$$

Тогда $P\left(\left\{\omega \,\middle|\, \xi_n(\omega) \xrightarrow[n\to\infty]{} 0\right\}\right) = P\left((0;1]\right) = 1$, так как $\lim\limits_{n\to\infty} \xi_n(\omega) \neq 0$ только при $\omega = 0$. Таким образом, $\xi_n \xrightarrow[n\to\infty]{\text{п.н.}} \xi$. При этом $E|\xi_n|^2 = n \cdot P(\xi_n = \sqrt{n}) = n \cdot P\left(\left[0;\frac{1}{n}\right]\right) = 1$, следовательно, последовательность не сходится к $\xi = 0$ в среднеквадратичном. ■

10.2. Сходимость последовательностей случайных величин по распределению

Рассмотрим еще один вид сходимости последовательностей случайных величин, который принципиально отличается от рассмотренных выше.

Определение 10.2.1. Последовательность случайных величин $\{\xi_n\}$ сходится ***по распределению*** к случайной величине ξ, если сходимость $F_{\xi_n(x)} \xrightarrow[n\to\infty]{} F_\xi(x)$ имеет место в точках непрерывности $F_\xi(x)$. При этом пишут $\xi_n \xrightarrow[n\to\infty]{d} \xi$.

R Если в определении 10.2.1 убрать требование сходимости последовательности функций распределения $F_{\xi_n}(x)$ к $F_\xi(x)$ только в точках непрерывности $F_\xi(x)$, мы получим ситуацию, когда, например, последовательность $\xi_n \equiv -\dfrac{1}{n}$ не сходится по распределению к $\xi \equiv 0$. Действительно, имеем:

$$F_{\xi_n}(x) = \begin{cases} 1\,, & x > -\frac{1}{n}\,, \\ 0\,, & x \le -\frac{1}{n}\,, \end{cases} \quad F_\xi(x) = \begin{cases} 1\,, & x > 0\,, \\ 0\,, & x \le 0\,. \end{cases}$$

В результате $F_{\xi_n}(0) \nrightarrow F_\xi(0)$ при $n \to \infty$.

R В отличие от определенных выше сходимостей почти наверное, в среднеквадратичном и по вероятности, для сходимости по распределению не требуется, чтобы все случайные величины ξ_n и ξ были определены на одном вероятностном пространстве.

R Если $\{F_n\}$ — последовательность функций распределения, $F(x)$ — функция распределения и $F_n(x) \xrightarrow[n\to\infty]{} F(x)$ в точках непрерывности $F(x)$, говорят, что последовательность распределелений вероятностей с функциями распределения F_n ***слабо сходится*** к распределению вероятностей с функцией распределения F, и пишут $F_n \Rightarrow F$. Таким образом, $\xi_n \xrightarrow[n\to\infty]{d} \xi \;\Leftrightarrow\; F_{\xi_n} \Rightarrow F_\xi$.

R Если все случайные величины ξ_n и ξ определены на одном вероятностном пространстве, сходимость по распределению для них можно сравнивать с другими видами сходимости, например, с сходимостью по вероятности. Можно доказать, что сходимость по вероятности влечет за собой сходимость по распределению. При этом оказывается, что сходимость по распределению, вообще говоря, не влечет за собой сходимость по вероятности. Рассмотрим, например, на вероятностном пространстве $(\Omega, \mathcal{F}, P)$, где $\Omega = [0;1]$, $\mathcal{F}$ — совокупность измеримых по Лебегу подмножеств этого отрезка, а P — мера Лебега на прямой, последовательность случайных величин $\{\xi_n\}$, где

$$\xi_{2n-1}(\omega) = \begin{cases} 1\,, & \omega \in \left(\frac{1}{2};1\right]\,, \\ -1\,, & \omega \in \left[0;\frac{1}{2}\right]\,, \end{cases} \quad \xi_{2n}(\omega) = \begin{cases} -1\,, & \omega \in \left(\frac{1}{2};1\right]\,, \\ 1\,, & \omega \in \left[0;\frac{1}{2}\right]\,. \end{cases}$$

Для этой последовательности имеют место сходимости $\xi_n \xrightarrow[n\to\infty]{d} \xi_1$ и $\xi_n \xrightarrow[n\to\infty]{d} \xi_2$, при этом по вероятности эта последовательность ни к ξ_1, ни к ξ_2 не сходится.

Таким образом, сходимость случайных величин по распределению — более слабый вид сходимости, чем сходимость по вероятности.

Тем не менее, справедливо следующее предложение, которое мы приведем без доказательства.

Предложение 10.2.1. Пусть $\{\xi_n\}$ — последовательность случайных величин, определенных на одном вероятностном пространстве. Тогда

$$\xi_n \xrightarrow[n\to\infty]{d} a = \mathrm{Const} \;\Rightarrow\; \xi_n \xrightarrow[n\to\infty]{P} a\,.$$

10.3. Закон больших чисел

Теорема 10.3.1. — Слабый закон больших чисел. Пусть $\{\xi_n\}$ — последовательность независимых случайных величин, определенных на одном вероятностном пространстве $(\Omega, \mathcal{F}, P)$, таких, что $E\xi_n = a$, $D\xi_n = \sigma^2$ для любого $n \in \mathbb{N}$. Тогда

$$\eta_n = \frac{1}{n}\sum_{k=1}^{n}\xi_k \xrightarrow[n\to\infty]{P} a\,.$$

Доказательство. Пользуясь линейностью математического ожидания, получаем:

$$E\eta_n = E\left(\frac{1}{n}\sum_{k=1}^{n}\xi_k\right) = \frac{1}{n}\sum_{k=1}^{n}E\xi_k = a\,.$$

Из независимости случайных величин ξ_k и свойств дисперсии следует:

$$D\eta_n = D\left(\frac{1}{n}\sum_{k=1}^{n}\xi_k\right) = \frac{1}{n^2}\sum_{k=1}^{n}D\xi_k = \frac{\sigma^2}{n}\,.$$

В результате, пользуясь неравенством Чебышева, получаем:

$$\forall\,\varepsilon > 0 \quad P\left(|\eta_n - a| > \varepsilon\right) \le \frac{D\eta_n}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \xrightarrow[n\to\infty]{} 0\,.$$

■

Доказанная теорема говорит о том, что если выполнены ее условия, при сложении независимых случайных слагаемых происходит «взаимопогашение» случайностей, в результате которого при усреднении этих сумм и неограниченном росте количества слагаемых в пределе получается детерминированная величина, т. е. случайности нивелируются. Сходимость по вероятности означает, что отклонения средних арифметических $\frac{1}{n}\sum_{k=1}^{n}\xi_k$ от a при больших n маловероятны.

Справедливо и более сильное утверждение, которое мы приведем без доказательства.

Теорема 10.3.2. — Усиленный закон больших чисел. Пусть $\{\xi_n\}$ — последовательность независимых случайных величин, определенных на одном вероятностном пространстве $(\Omega, \mathcal{F}, P)$, таких, что $E\xi_n = a$, $D\xi_n = \sigma^2$ для любого $n \in N$. Тогда

$$\eta_n = \frac{1}{n}\sum_{k=1}^{n}\xi_k \xrightarrow[n\to\infty]{\text{п.н.}} a\,.$$

Следствие 10.3.3. Рассмотрим серию испытаний Бернулли с вероятностью успеха p. Пусть ξ_k — индикатор успеха в k-м испытании, $\sum_{k=1}^{n}\xi_k$ — число успехов в серии из n испытаний Бернулли. Тогда частота наступления успеха в n испытаниях $\frac{1}{n}\sum_{k=1}^{n}\xi_k$ с вероятностью 1 сходится к p при $n \to \infty$.

Это утверждение сразу получается из усиленного закона больших чисел, поскольку $E\xi_k = p$. Таким образом, закон больших чисел обосновывает эмпирически наблюдаемое явление, благодаря которому понятие вероятности исторически возникло как мера частоты наступления случайного события в серии повторяющихся одинаковых независимых случайных экспериментов.

Сходимость (в более слабом смысле) средних арифметических независимых случайных величин к их общему математическому ожиданию имеет место и при более слабых условиях. Для получения этого результата рассмотрим следующую теорему, которая служит инструментом для доказательства многих предельных теорем теории вероятностей.

10.4. Теорема непрерывности

Эта теорема устанавливает связь между слабой сходимостью распределений вероятностей (сходимостью случайных величин по распределению — см. замечание к определению 10.2.1 на стр. 154) и сходимостью их характеристических функций.

Теорема 10.4.1. — Теорема непрерывности. Пусть $\{\varphi_n\}$ — последовательность характеристических функций, соответствующих распределениям вероятностей с функциями распределения $F_n(x)$.

1. Если $F_n \Rightarrow F$ и F — функция распределения, то $\varphi_n(t) \xrightarrow[n\to\infty]{\mathbb{R}} \varphi(t)$, где $\varphi(t)$ — характеристическая функция распределения вероятностей, соответствующего F;

2. Если $\varphi_n(t) \xrightarrow[n\to\infty]{\mathbb{R}} \varphi(t)$ и $\varphi(t)$ непрерывна в точке $t = 0$, то $\varphi(t)$ — характеристическая функция некоторого распределения вероятностей и $F_n \Rightarrow F$, где F — функция распределения вероятностей, соответствующего характеристической функции φ.

Теорема называется теоремой непрерывности из-за условия непрерывности в точке $t = 0$, которое накладывается на функцию $\varphi(t)$ во втором утверждении и которое обеспечивает то, что эта функция оказывается характеристической для некоторого распределения вероятностей. Мы приводим эту теорему без доказательства.

Рассмотрим несколько предельных теорем, которые доказываются с помощью теоремы непрерывности.

10.5. Закон больших чисел Хинчина

Теорема 10.5.1. — Закон больших чисел Хинчина. Пусть $\{\xi_n\}$ — последовательность независимых случайных величин, $E\xi_n = a \in \mathbb{R}$ для любого $n \in \mathbb{N}$. Тогда

$$\eta_n = \frac{1}{n}\sum_{k=1}^{n}\xi_k \xrightarrow[n\to\infty]{P} a\,.$$

Доказательство. В силу предложения 10.2.1 достаточно доказать, что имеет место сходимость $\eta_n = \dfrac{1}{n}\sum\limits_{k=1}^{n} \xi_k \xrightarrow[n\to\infty]{d} a$.

Рассмотрим характеристические функции случайных величин ξ_n. По формуле Тейлора–Маклорена с остаточным членом в форме Пеано получим:

$$\varphi_{\xi_n}(t) = \varphi_{\xi_n}(0) + \varphi'_{\xi_n}(0)t + o(t) = 1 + iat + o(t)\,, \quad t \to 0\,.$$

Здесь мы воспользовались свойствами (φ-I) и (φ-III) характеристических функций. Пользуясь свойствами (φ-II) и (φ-V) для характеристической функции случайной величины η_n, получим разложение

$$\varphi_{\eta_n}(t) = \varphi_{\frac{1}{n}\sum\limits_{k=1}^{n}\xi_k}(t) = \varphi_{\sum\limits_{k=1}^{n}\xi_k}\left(\frac{t}{n}\right) = \prod_{k=1}^{n}\varphi_{\xi_k}\left(\frac{t}{n}\right) =$$
$$= \left(1 + \frac{iat}{n} + o\left(\frac{t}{n}\right)\right)^n \xrightarrow[n\to\infty]{} e^{iat} = Ee^{iat} = \varphi_a(t)\,.$$

По теореме непрерывности отсюда следует сходимость η_n по распределению к a. ■

R Заметим, что в отличие от слабого закона больших чисел, закон больших чисел Хинчина не требует ни существования дисперсий случайных величин ξ_n, ни их равенства между собой.

R Условие конечности математического ожидания случайных величин ξ_n существенно для закона больших чисел. Если от него отказаться, явление «нивелирования случайностей», о котором говорит закон больших чисел, перестает иметь место. Действительно, пусть, например, случайные величины ξ_n независимы и $\xi_n \sim C(0,1)$, т. е. их плотности распределения имеют вид $f_{\xi_n}(x) = \dfrac{1}{\pi(1+x^2)}$. Тогда

$$\varphi_{\xi_k}(t) = Ee^{i\xi_k t} = \frac{1}{\pi}\int\limits_{-\infty}^{+\infty} \frac{e^{ixt}}{(1+x^2)}\,dx = e^{-|t|}\,.$$

Последнее равенство можно получить по известным формулам теории вычетов:

Если $f(z)$ аналитична в $\mathbb{C}$ всюду, за исключением конечного числа изолированных особых точек z_k, не лежащих на действительной оси, и $\max\limits_{|z|=R} |f(z)| \xrightarrow[R\to+\infty]{} 0$, то

$$\int\limits_{-\infty}^{+\infty} e^{ixt} f(x)\,dx = \begin{cases} 2\pi i \sum\limits_{\operatorname{Im} z_k>0} \operatorname*{res}\limits_{z=z_k} e^{izt} f(z)\,, & t>0\,,\\ 2\pi i \sum\limits_{\operatorname{Im} z_k<0} \operatorname*{res}\limits_{z_k} e^{izt} f(z)\,, & t<0\,. \end{cases}$$

Как уже отмечалось, случайные величины, распределенные по закону Коши, не имеют моментов, в частности, $E|\xi_k| = \infty$ (см. замечание на стр. 119). Это, в силу свойства (φ-III) характеристических функций, согласуется с тем, что $\varphi_{\xi_k}(t)$ не дифференцируема в точке $t = 0$. Пользуясь свойствами (φ-II) и (φ-V) характеристических функций, получаем:

$$\varphi_{\frac{1}{n}\sum\limits_{k=1}^{n}\xi_k}(t) = \left(\varphi_{\xi_k}\left(\frac{t}{n}\right)\right)^n = e^{-|t|}\,.$$

Отсюда следует, что в данном случае $\dfrac{1}{n}\sum\limits_{k=1}^{n}\xi_k \sim C(0,1)$. Таким образом, сложение стремящегося к бесконечности количества одинаковых независимых случайных величин с последующим усреднением суммы по количеству слагаемых в данном случае не приводит к тому, что «случайность исчезает».

10.6. Центральная предельная теорема

Теоремы Муавра–Лапласа, рассмотренные ранее в главе 4, характеризуют поведение числа успехов в схеме Бернулли при стремлении числа испытаний к бесконечности. Используя введенные выше понятия, утверждение интегральной теоремы Муавра–Лапласа можно сформулировать в терминах сходимости последовательности случайных величин.

Пусть ξ_k — индикатор успеха в k-м испытании серии Бернулли с вероятностью успеха p, т. е. ξ_k принимает значение 1, если k-е испытание закончилось успехом, и 0, если неудачей. Тогда $\theta_n = \sum\limits_{k=1}^{n}\xi_k$ — число успехов в серии из n испытаний Бернулли и $\theta_n \sim B(n,p)$. При этом $E\xi_k = p$, $D\xi_k = E\xi_k^2 - (E\xi_k)^2 = p - p^2 = pq$, следовательно,

$E\theta_n = np$, $D_n = npq$. Интегральная теорема Муавра–Лапласа утверждает, что

$$\frac{\theta_n - E\theta_n}{\sqrt{\theta_n}} = \frac{\sum\limits_{k=1}^{n} \xi_k - E\left(\sum\limits_{k=1}^{n} \xi_k\right)}{\sqrt{D\left(\sum\limits_{k=1}^{n} \xi_k\right)}} \xrightarrow[n\to\infty]{d} \eta \sim N(0,1)\,. \tag{10.2}$$

Принято говорить, что ***для последовательности случайных величин $\{\xi_n\}$ справедлива Центральная Предельная Теорема***, если имеет место сходимость (10.2). Таким образом, интегральная теорема Муавра–Лапласа утверждает, что для последовательности независимых случайных величин $\{\xi_n\}$, где ξ_n — индикатор успеха в n-м испытании Бернулли, выполняется Центральная Предельная Теорема.

Явление, которое описывает Центральная Предельная Теорема, подобно Закону Больших Чисел, рассмотренному выше. Его суть также состоит в том, что при суммировании большого числа независимых случайных величин их индивидуальные характеристики нивелируются, при этом центрирование таких сумм и нормирование на их среднеквадратическое отклонение делает их близкими по распределению к стандартной гауссовской случайной величине.

Оказывается, это явление происходит не только в частном случае, описываемом теоремой Муавра–Лапласа. Оно имеет место и тогда, когда вместо индикаторов — случайных величин, принимающих значения 1 и 0 — берут произвольные случайные величины, при этом даже не обязательно одинаково распределенные.

Теорема 10.6.1. — Центральная Предельная Теорема I. Пусть $\{\xi_n\}$ — последовательность независимых случайных величин, таких, что $E\xi_n = a$ и $D\xi_n = \sigma^2$ для любого $n \in \mathbb{N}$. Тогда

$$\eta_n = \frac{\sum\limits_{k=1}^{n} \xi_k - na}{\sigma\sqrt{n}} \xrightarrow[n\to\infty]{d} \eta \sim N(0,1)\,. \tag{10.3}$$

Доказательство. По формуле Тейлора–Маклорена для характеристической функции центрированной случайной величины $\xi_n - a$ справедливо разложение:

$$
\begin{aligned}
\varphi_{\xi_n-a}(t) &= \varphi_{\xi_n-a}(0) + \varphi'_{\xi_n-a}(0)\,t + \frac{1}{2}\varphi''_{\xi_n-a}(0)\,t^2 + o(t^2) = \\
&= 1 - \frac{\sigma^2 t^2}{2} + o(t^2)\,, \quad t \to 0\,.
\end{aligned} \tag{10.4}
$$

Здесь мы воспользовались равенствами

$$
\varphi'_{\xi_n-a}(0) = iE(\xi_n - a) = 0\,, \quad \varphi''_{\xi_n-a}(0) = -E(\xi_n - a)^2 = -\sigma^2\,.
$$

Для характеристической функции случайной величины

$$
\eta_n = \frac{\sum\limits_{k=1}^{n} \xi_k - na}{\sigma\sqrt{n}} = \frac{1}{\sigma\sqrt{n}} \sum_{k=1}^{n} (\xi_k - a)\,,
$$

пользуясь разложением (10.4), получаем:

$$
\begin{aligned}
\varphi_{\eta_n}(t) &= \varphi_{\sum\limits_{k=1}^{n}(\xi_k-a)}\left(\frac{t}{\sigma\sqrt{n}}\right) = \prod_{k=1}^{n} \varphi_{\xi_k-a}\left(\frac{t}{\sigma\sqrt{n}}\right) = \\
&= \left(1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right)^n \xrightarrow[n\to\infty]{} e^{-\frac{t^2}{2}} = \varphi_{\eta}(t)\,,
\end{aligned}
$$

где $\eta \sim N(0,1)$ (см. пример 9.2.2). Асимптотическая формула (10.4) работает, потому что $\dfrac{t}{\sigma\sqrt{n}} \to 0$ при фиксированном t и $n \to \infty$. По теореме непрерывности отсюда следует сходимость (10.3). ■

Центральную Предельную Теорему можно доказать и при еще более слабых условиях, а именно отказавшись от требования того, чтобы ξ_n имели одинаковые математические ожидания и дисперсии.

Теорема 10.6.2. — Центральная Предельная Теорема Ляпунова. Пусть $\{\xi_k\}$ — последовательность независимых случайных величин с математическими ожиданиями $E\xi_k = a_k$, дисперсиями $D\xi_k = b_k^2$ и абсолютными центральными моментами третьего порядка $E|\xi_k - a_k|^3 = c_k^3$. Пусть $A_n = \sum\limits_{k=1}^{n} a_k$, $B_n^2 = \sum\limits_{k=1}^{n} b_k^2$, $C_n^3 = \sum\limits_{k=1}^{n} c_k^3$, причем выполнено условие

$$
\frac{C_n}{B_n} \xrightarrow[n\to\infty]{} 0\,. \tag{10.5}
$$

Тогда

$$\eta_n = \frac{\sum\limits_{k=1}^{n} \xi_k - A_n}{B_n} \xrightarrow[n\to\infty]{d} \eta \sim N(0,1)\,. \tag{10.6}$$

Доказательство. Для характеристической функции центрированной случайной величины $\xi_k - a_k$, по формуле Тейлора–Маклорена второго порядка с остаточным членом в форме Лагранжа, получаем разложение:

$$\varphi_{\xi_k-a_k}(t) = \varphi_{\xi_k-a_k}(0) + \varphi'_{\xi_k-a_k}(0)\,t + \frac{1}{2}\varphi''_{\xi_k-a_k}(0)\,t^2 + \frac{1}{3!}\varphi'''_{\xi_k-a_k}(\theta t)\,t^3 = \\ = 1 - \frac{b_k^2 t^2}{2} + \frac{R_k}{6}t^3\,,$$

где $\theta \in (0;1)$ — некоторое число,

$$R_k := \varphi'''_{\xi_k-a_k}(\theta t) = E\left(-i(\xi_k - a_k)^3 e^{i\theta(\xi_k-a_k)t}\right)\,.$$

Здесь мы также использовали равенства

$$\varphi'_{\xi_k-a_k}(0) = iE(\xi_k - a_k) = 0\,, \quad \varphi''_{\xi_k-a_k}(0) = -E(\xi_k - a_k)^2 = -b_k^2\,.$$

Пользуясь полученным разложением для характеристической функции случайной величины $\eta_n = \dfrac{\sum\limits_{k=1}^{n} \xi_k - A_n}{B_n} = \dfrac{1}{B_n}\sum\limits_{k=1}^{n}(\xi_k - a_k)$, получаем:

$$\varphi_{\eta_n}(t) = \varphi_{\sum\limits_{k=1}^{n}(\xi_k-a_k)}\left(\frac{t}{B_n}\right) = \prod_{k=1}^{n}\varphi_{\xi_k-a_k}\left(\frac{t}{B_n}\right) = \\ = \prod_{k=1}^{n}\left(1 - \frac{b_k^2t^2}{2B_n^2} + \frac{R_k t^3}{6B_n^3}\right) = \exp\left[\sum_{k=1}^{n}\ln\left(1 - \frac{b_k^2t^2}{2B_n^2} + O\left(\frac{c_k^3}{B_n^3}\right)\right)\right] = \\ = \exp\left[\sum_{k=1}^{n}\left(-\frac{b_k^2t^2}{2B_n^2} + O\left(\frac{c_k^3}{B_n^3}\right)\right)\right] = \\ = \exp\left[-\frac{t^2}{2} + O\left(\frac{C_n^3}{B_n^3}\right)\right] \xrightarrow[n\to\infty]{} e^{-\frac{t^2}{2}} = \varphi_\eta(t)\,.$$

Здесь мы воспользовались очевидной оценкой $|R_k| \le E|\xi_k - a_k|^3 = c_k^3$, разложением $\ln(1+x) = x + o(x)\,, x \to 0$ логарифма по формуле

Тейлора–Маклорена и условием (10.5). Из доказанной сходимости характеристических функций, по теореме непрерывности, следует сходимость (10.6). ■

Глава 11. Основы математической статистики

Математическая статистика изучает математический аппарат, использующийся для обработки результатов статистических экспериментов. Цель статистического эксперимента — получение информации о свойствах случайного явления. Основным понятием математической статистики является понятие выборки. Выборка представляет собой математическую модель статистического эксперимента, объектом изучения в котором является некоторая случайная величина.

Определение 11.0.1. Пусть ξ — случайная величина, распределение вероятностей P_ξ которой задано функцией распределения $F_\xi(x)$. Случайный вектор $\vec{\xi_n} = (\xi_1, \xi_2, ..., \xi_n)^T$, где случайные величины ξ_i независимы и $F_{\xi_i}(x) = F_\xi(x)\,, i = 1, 2, \ldots, n$, называют ***выборкой объема n, отвечающей случайной величине ξ*** (или выборкой объема n из генеральной совокупности с распределением P_ξ[a]).

Если в результате проведения статистического эксперимента компоненты выборки приняли значения $\xi_1 = x_1, \ldots, \xi_n = x_n$, то вектор $\overline{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ называют ***реализацией выборки***. Пространство $\mathbb{R}^n$ при этом называют ***выборочным пространством***.

[a]Использование этой терминологии — дань определенной исторической традиции.

Проведение статистического эксперимента, результатом которого является некоторая реализация $\overline{x}$ выборки $\vec{\xi_n}$ объема n, можно считать n-кратным независимым повторением одного и того же случайного эксперимента, в котором наблюдается случайная величина ξ, а i-я компонента выборки ξ_i моделирует результат i-го повторения.

Основная задача математической статистики — построение оценок параметров распределения P_ξ по выборке $\vec{\xi_n}$. Решается она с помощью ***статистик***.

Определение 11.0.2. Пусть $g_n : \mathbb{R}^n \to \mathbb{R}$ — измеримая по Борелю функция. Случайная величина $g_n\big(\vec{\xi_n}\big)$ называется ***статистикой***.

Пусть θ — некоторый параметр распределения P_ξ. Для его оценки функцию g_n подбирают так, чтобы $g_n(\vec{\xi_n}) \approx \theta$. Приближение считается хорошим, если статистика $g_n(\vec{\xi_n})$ при $n \to \infty$ в некотором смысле сходится к θ.

Начнем с оценок таких параметров P_ξ, как функция распределения $F_\xi(x)$ и плотность распределения $f_\xi(x)$.

11.1. Эмпирическая функция распределения

Для оценки $F_\xi(x)$ используется ***эмпирическая функция распределения $\boldsymbol{F_n(x)}$***, которая определяется равенством

$$F_n(x) = F_n(x)[\vec{\xi_n}] = \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{(-\infty;x)}(\xi_k)\,. \tag{11.1}$$

R Эмпирическая функция распределения представляет собой функцию от выборки, а значит является случайной величиной. В результате при фиксированном n каждое повторение статистического эксперимента дает некоторую реализацию выборки и соответствующую этой реализации (каждый раз новую) эмпирическую функцию распределения ξ.

Случайная величина $\eta_k = \mathbf{1}_{(-\infty;x)}(\xi_k)$ принимает значение 1 с вероятностью, равной $P\big(\xi_k \in (-\infty;x)\big) = F_\xi(x)$, и 0 с вероятностью

$1 - F_\xi(x)$, поэтому $E\eta_k = F_\xi(x)$. Поскольку η_k независимы, по усиленному закону больших чисел (Теорема 10.3.2), имеет место сходимость

$$F_n(x) = \frac{1}{n}\sum_{k=1}^{n} \eta_k \xrightarrow[n\to\infty]{\text{п.н.}} E\eta_k = F_\xi(x)\,.$$

В результате эмпирическая функция распределения $F_n(x)$ при достаточно больших n хорошо приближает $F_\xi(x)$. На самом деле это приближение является равномерным на $\mathbb{R}$. Это утверждает следующая теорема, которую мы приведем без доказательства.

Теорема 11.1.1. — Гливенко. Имеет место сходимость

$$\sup_{x\in\mathbb{R}} |F_n(x) - F_\xi(x)| \xrightarrow[n\to\infty]{\text{п.н.}} 0\,.$$

11.2. Эмпирическая плотность распределения

Для оценки $f_\xi(x)$ используется ***эмпирическая плотность распределения*** $\boldsymbol{f_{nh}(x)}$, которая определяется равенством

$$f_{nh}(x) = f_{nh}(x)[\vec{\xi}_n] = \frac{1}{nh}\sum_{k=1}^{n} \mathbf{1}_{[x_h; x_h+h)}(\xi_k)\,, \tag{11.2}$$

где $h > 0$, $x_h = \left[\dfrac{x}{h}\right] h$.

Параметр h задает разбиение $\mathbb{R}$ точками вида kh, где $k \in \mathbb{Z}$. Любое $x \in \mathbb{R}$ попадает в один и только один из полуинтервалов вида $\big[kh; (k+1)h\big)$. Действительно, по определению целой части числа

$$\left[\frac{x}{h}\right] \le \frac{x}{h} < \left[\frac{x}{h}\right] + 1 \;\Rightarrow\; \left[\frac{x}{h}\right] h \le x < \left[\frac{x}{h}\right] h + h\,,$$

т. е. $x \in \big[kh; (k+1)h\big)$, где $k = \left[\dfrac{x}{h}\right]$.

Величина $\dfrac{1}{n}\displaystyle\sum_{k=1}^{n} \mathbf{1}_{[x_h; x_h+h)}(\xi_k)$ равна частоте попадания значений компонент выборки $\vec{\xi}_n$ в промежуток $[x_h; x_h+h)$. По следствию 10.3.3

из усиленного закона больших чисел при больших n эта частота оказывается близкой к вероятности $P\big(\xi \in [x_h; x_h + h)\big)$. При малых значениях h эта вероятность, в свою очередь, близка к $f_\xi(x_h)h$, таким образом, получается

$$f_{nh}(x_h)h = \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{[x_h;x_h+h)}(\xi_k) \approx f_\xi(x_h)h\,.$$

Более точно справедлива следующая теорема.

> Теорема 11.2.1. В точках непрерывности $f_\xi(x)$ имеет место сходимость
>
> $$f_{nh}(x) \xrightarrow[\substack{n\to\infty \\ h\to 0 \\ nh\to\infty}]{P} f_\xi(x)\,.$$

Доказательство теоремы заменим следующим нестрогим обоснованием. Пусть $\eta_k = \dfrac{1}{h}\cdot \mathbf{1}_{[x_h;x_h+h)}(\xi_k)$. Тогда $f_{nh}(x) = \dfrac{1}{n}\sum\limits_{k=1}^{n}\eta_k$ и, так как случайные величины η_k независимы, по слабому закону больших чисел получаем $f_{nh}(x) \xrightarrow[n\to\infty]{P} E\eta_k$. При этом

$$E\eta_k = \frac{1}{h}\cdot P(x_h \le \xi_k < x_h + h) = \frac{1}{h}\cdot \int\limits_{x_h}^{x_h+h} f_\xi(x)\,dx =$$
$$= \frac{1}{h}\cdot f_\xi(x_h + \theta h)\cdot h = f_\xi(x_h+\theta h)\,,$$

для некоторого $\theta \in (0;1)$ по интегральной теореме о среднем. Поскольку $x_h + \theta h \to x$ при $h \to 0$, в силу непрерывности f_ξ в точке x получаем $f_\xi(x_h+\theta h) \xrightarrow[h\to 0]{} f_\xi(x)$.

11.3. Точечные оценки моментов распределения

Пусть θ — некоторый параметр распределения вероятностей P_ξ, а $g_n\big(\vec{\xi}_n\big)$ — некоторая статистика, используемая для его оценки. Такие оценки называют ***точечными***. Рассмотрим понятия, использующиеся для характеристики качества точечных статистических оценок.

Определение 11.3.1. Оценка $g_n\big(\vec{\xi_n}\big)$ параметра θ называется ***состоятельной***, если имеет место сходимость

$$g_n\big(\vec{\xi_n}\big) \xrightarrow[n\to\infty]{P} \theta\,.$$

Определение 11.3.2. Оценка $g_n\big(\vec{\xi_n}\big)$ параметра θ называется ***несмещенной***, если $Eg_n\big(\vec{\xi_n}\big) = \theta$.

Определение 11.3.3. Оценка $g_n\big(\vec{\xi_n}\big)$ параметра θ называется ***инвариантной относительно сдвига***, если

$$\forall\, c \in \mathbb{R} \quad g_n(\xi_1 + c, \xi_2 + c, \ldots, \xi_n + c) = g_n(\xi_1, \xi_2, \ldots, \xi_n)\,.$$

Наиболее часто используют следующие точечные оценки моментов распределения вероятностей P_ξ:

Выборочное среднее:

$$\overline{\xi_n} := \frac{1}{n}\sum_{k=1}^{n} \xi_k\,;$$

Выборочный момент i-го порядка:

$$\overline{\xi_n}^{(i)} := \frac{1}{n}\sum_{k=1}^{n} \xi_k^i\,;$$

Выборочный центральный момент i-го порядка:

$$S_n^{(i)} := \frac{1}{n}\sum_{k=1}^{n} \big(\xi_k - \overline{\xi_n}\big)^i\,.$$

Предложение 11.3.1. Выборочное среднее является состоятельной и несмещенной оценкой $E\xi$.

Доказательство. Пользуясь линейностью математического ожидания, получаем:

$$E\overline{\xi_n} = E\left(\frac{1}{n}\sum_{k=1}^{n}\xi_k\right) = \frac{1}{n}\sum_{k=1}^{n}E\xi_k = \frac{1}{n}\sum_{k=1}^{n}E\xi = E\xi\,.$$

Состоятельность выборочного среднего сразу следует из слабого закона больших чисел (Теорема 10.3.1). ■

Предложение 11.3.2. Выборочный момент i-го порядка является состоятельной и несмещенной оценкой $E\xi^i$.

Доказательство. Также как в доказательстве предыдущего предложения, получаем:

$$E\overline{\xi_n}^{(i)} = E\left(\frac{1}{n}\sum_{k=1}^{n}\xi_k^i\right) = \frac{1}{n}\sum_{k=1}^{n}E\xi_k^i = \frac{1}{n}\sum_{k=1}^{n}E\xi^i = E\xi^i\,.$$

Состоятельность получаем, применяя закон больших чисел к последовательности $\{\xi_k^i\}$. ■

Без подробного доказательства приведем следующее предложение.

Предложение 11.3.3. Выборочный центральный момент 2-го порядка является состоятельной, инвариантной относительно сдвига оценкой $D\xi$.

Отметим только, что инвариантность этой оценки относительно сдвига легко проверяется с помощью свойств математического ожидания, так как

$$\overline{\xi_n + c} := \frac{1}{n}\sum_{k=1}^{n}(\xi_k + c) = \overline{\xi_n} + c$$

и, следовательно,

$$\frac{1}{n}\sum_{k=1}^{n}\left(\xi_k + c - \overline{\xi_n + c}\right)^2 = \frac{1}{n}\sum_{k=1}^{n}\left(\xi_k - \overline{\xi_n}\right)^2 = S_n^{(2)}\,.$$

(R) Выборочное среднее и выборочный момент i-го порядка, очевидно, не являются оценками, инвариантными относительно сдвига.

(R) Выборочный центральный момент i-го порядка не является несмещенной оценкой $D\xi$ центрального момента i-го порядка. В частности, ***выборочный центральный момент 2-го порядка не является***

несмещенной оценкой дисперсии. Покажем это. Действительно, пусть $E\xi_k = E\xi = a$, тогда

$$ES_n^{(2)} = E\left(\frac{1}{n}\sum_{k=1}^{n}\left(\xi_k - \frac{1}{n}\sum_{i=1}^{n}\xi_i\right)^2\right) =$$
$$= \frac{1}{n}\sum_{k=1}^{n} E\left(\xi_k - a - \frac{1}{n}\sum_{i=1}^{n}(\xi_i - a)\right)^2 .$$

Возводя в квадрат сумму $n+1$ слагаемых, получаем

$$ES_n^{(2)} = \frac{1}{n}\sum_{k=1}^{n} E\Bigg[(\xi_k - a)^2 + \sum_{i=1}^{n}\frac{(\xi_i - a)^2}{n^2} - \frac{2}{n}\sum_{i=1}^{n}(\xi_k - a)(\xi_i - a) +$$
$$+ \frac{2}{n^2}\sum_{1\le i<j\le n}(\xi_i - a)(\xi_j - a)\Bigg].$$

Пользуясь тем, что $E(\xi_i - a)^2 = E(\xi - a)^2 = D\xi$ для любого $i = 1, 2, ..., n$, и тем, что, в силу независимости компонент выборки, $E\left[(\xi_i - a)(\xi_j - a)\right] = E(\xi_i - a)\, E(\xi_j - a) = 0$ при $i \neq j$, получаем

$$ES_n^{(2)} = \frac{1}{n}\sum_{k=1}^{n}\left[D\xi + \frac{1}{n}D\xi - \frac{2}{n}D\xi\right] = \left(1 - \frac{1}{n}\right) D\xi\,.$$

Поэтому для дисперсии чаще используют другую оценку — выборочную дисперсию.

Выборочная дисперсия:

$$S_n^2 := \frac{n}{n-1}S_n^{(2)} = \frac{1}{n-1}\sum_{k=1}^{n}\left(\xi_k - \overline{\xi_n}\right)^2 .$$

Из доказанного выше следует ее несмещенность. Остальные утверждения следующего предложения примем без доказательства.

Предложение 11.3.4. Выборочная дисперсия является состоятельной, несмещенной, инвариантной относительно сдвига оценкой $D\xi$.

11.4. Интервальные оценки моментов распределения

Точечные оценки параметров распределения P_ξ обладают существенным недостатком. По своей природе они являются случайными величинами, функциями от выборки. Поэтому, когда проводится статистический эксперимент, результатом которого является некоторая реализация этой выборки, оценка параметра, если она состоятельна, принимает некоторое числовое значение, о котором известно только то, что оно близко к истинному значению оцениваемого параметра. Чем больше объем выборки, тем меньше вероятность большого отклонения этого значения от истинного, однако неизвестно, насколько именно и в какую сторону оно отклоняется. Интервальные оценки позволяют в определенном смысле устранить этот недостаток.

Определение 11.4.1. Интервал $\left(\underline{\theta}\left(\vec{\xi}_n\right);\overline{\theta}\left(\vec{\xi}_n\right)\right)$ называется ***доверительным*** для оценки параметра θ, отвечающим доверительной вероятности α, если

$$P\left(\left(\underline{\theta}(\vec{\xi}_n);\overline{\theta}(\vec{\xi}_n)\right) \ni \theta\right) \geq \alpha\,.$$

Рассмотрим построение доверительного интервала для $E\xi$, основанное на использовании свойства асимптотической нормальности выборочного среднего.

Если статистика $g_n\left(\vec{\xi}_n\right)$ представляет собой состоятельную оценку некоторого параметра θ, то разность $g_n\left(\vec{\xi}_n\right) - \theta$ стремится к нулю по вероятности при $n \to \infty$. При этом домножение этой разности на подходящую бесконечно большую величину может привести к тому, что пределом произведения станет случайная величина, отличная от нуля.

Определение 11.4.2. Оценка $g_n\left(\vec{\xi}_n\right)$ параметра θ распределения вероятностей P_ξ называется ***асимптотически нормальной с дисперсией*** $\boldsymbol{\Delta^2}$, если

$$\frac{\sqrt{n}}{\Delta}\left(g_n\left(\vec{\xi}_n\right) - \theta\right) \xrightarrow[n\to\infty]{d} \eta \sim N(0,1)\,.$$

Предложение 11.4.1. Выборочное среднее является асимптотически нормальной оценкой $E\xi$ с дисперсией $\sigma^2 = D\xi$.

Доказательство. Пусть $\vec{\xi}_n = (\xi_1, \ldots, \xi_n)^T$ — выборка, отвечающая случайной величине ξ. Пусть $E\xi_k = E\xi = a\,,\ D\xi_k = D\xi = \sigma^2$. Случайные величины ξ_k независимы, поэтому в силу центральной предельной теоремы (Теорема 10.6.1), имеем:

$$\frac{\sqrt{n}}{\sigma}\left(\overline{\xi_n}-a\right)=\frac{\sum\limits_{k=1}^{n}\xi_k-na}{\sigma\sqrt{n}}\xrightarrow[n\to\infty]{d}\eta\sim N(0,1)\,. \qquad \blacksquare$$

Из предложения 11.4.1 следует, что при достаточно больших значениях n можно считать, что $\dfrac{\sqrt{n}}{\sigma}\left(\overline{\xi_n}-a\right)\approx N(0,1)$. В результате имеем:

$$P\left(\left|\overline{\xi_n}-a\right|<\varepsilon\right)=P\left(\frac{\sqrt{n}}{\sigma}\left|\overline{\xi_n}-a\right|<\frac{\varepsilon\sqrt{n}}{\sigma}\right)\approx\Phi_0(x)\Big|_{-\frac{\varepsilon\sqrt{n}}{\sigma}}^{\frac{\varepsilon\sqrt{n}}{\sigma}}$$
$$=\Phi_0\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right)-\left(1-\Phi_0\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right)\right)=2\Phi_0\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right)-1\,,$$

где $\Phi_0(x)=\dfrac{1}{\sqrt{2\pi}}\displaystyle\int\limits_0^x e^{-\frac{t^2}{2}}dt$ — функция распределения стандартного нормального распределения вероятностей. Потребуем

$$2\Phi_0\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right)-1=\alpha\,.$$

Выражая из полученного равенства ε, получаем

$$\varepsilon=\frac{\sigma}{\sqrt{n}}\Phi_0^{-1}\left(\frac{\alpha+1}{2}\right)\,.$$

Таким образом, границы доверительного интервала для $E\xi$, отвечающего доверительной вероятности α, имеют вид:

$$\begin{aligned}\underline{\theta}\left(\vec{\xi_n}\right)&=\overline{\xi_n}-\frac{\sigma}{\sqrt{n}}\Phi_0^{-1}\left(\frac{\alpha+1}{2}\right)\,,\\ \overline{\theta}\left(\vec{\xi_n}\right)&=\overline{\xi_n}+\frac{\sigma}{\sqrt{n}}\Phi_0^{-1}\left(\frac{\alpha+1}{2}\right)\,.\end{aligned}\qquad(11.3)$$

R Из формул (11.3) видно, что так как $\Phi_0^{-1}(x)$ строго возрастает, увеличение доверительной вероятности приводит к расширению построенного доверительного интервала. Это естественно, так как увеличение доверительной вероятности означает большую надежность прогноза для оцениваемого математического ожидания. Кроме того, хорошо видно, что увеличение объема выборки приводит к сужению доверительного интервала, что тоже совершенно естественно, так как увеличение объема выборки приводит к увеличению количества информации о свойствах исследуемой случайной величины, а значит к уточнению оценки.

Приложение Е содержит задания лабораторной работы «Оценки параметров распределения вероятностей по выборке» (раздел Е.1), которую предлагается выполнить в вычислительной среде *MatLab*.

Глава 12. Проверка статистических гипотез

В этой главе мы рассмотрим простейшую задачу проверки статистической гипотезы о характере распределения вероятностей исследуемой случайной величины.

12.1. Постановка задачи и терминология

Пусть $\vec{\xi_n} = (\xi_1, \xi_2, \ldots, \xi_n)^T$ — выборка, соответствующая некоторой случайной величине ξ с функцией распределения $F_\xi(x)$. Рассмотрим две гипотезы о распределении вероятностей ξ:

$$
\begin{aligned}
&H_0 = \{F_\xi(x) \equiv F_0(x)\}\,,\\
&H_1 = \{F_\xi(x) \not\equiv F_0(x)\}\,.
\end{aligned}
$$

Гипотезу H_0 принято называть ***нуль-гипотезой***, а H_1 — ***альтернативной гипотезой***. Рассмотрим задачу о построении критерия согласия, позволяющего принять решение о выборе одной из двух гипотез.

Определение 12.1.1. Статистика $\eta\left(\vec{\xi_n}\right)$ называется ***критерием согласия***, если она принимает значения 1 или 0, при этом если $\eta\left(\vec{\xi_n}\right) = 1$, гипотеза H_0 принимается, а если $\eta\left(\vec{\xi_n}\right) = 0$, гипотеза H_0 отвергается и принимается H_1.

Применение критерия может привести к ошибке. Различают два вида ошибок. Если $\eta\big(\vec{\xi}_n\big) = 0$, т. е. критерий показывает, что следует отвергнуть нуль-гипотезу когда она верна, то говорят, что ***имеет место ошибка первого рода***. Вероятность ошибки первого рода, т. е. величину

$$\alpha = P\left(\eta\big(\vec{\xi}_n\big) = 0 \,\middle|\, H_0\right) ,$$

называют ***уровнем значимости критерия***.

Если $\eta\big(\vec{\xi}_n\big) = 1$, т. е. критерий показывает, что нуль-гипотеза должна быть принята, а при этом верна гипотеза H_1, говорят, что ***имеет место ошибка второго рода***. Вероятность ошибки второго рода будем обозначать через β:

$$\beta = P\left(\eta\big(\vec{\xi}_n\big) = 1 \,\middle|\, H_1\right) .$$

При заданном объеме выборки одновременно сделать α и β достаточно малыми невозможно, поэтому обычно задают α, т. е. желаемый уровень значимости критерия, и стараются минимизировать β при заданном α.

Определение 12.1.2. Критерий согласия $\eta\big(\vec{\xi}_n\big)$ называют ***состоятельным***, если $\beta \xrightarrow[n\to\infty]{} 0$ при фиксированном значении α.

Мы рассмотрим два критерия согласия — критерий Колмогорова и критерий Пирсона.

12.2. Критерий согласия Колмогорова

Критерий Колмогорова основан на теореме, которая описывает поведение при $n \to \infty$ величины $\sup\limits_{x\in\mathbb{R}} |F_n(x) - F_\xi(x)|$, где $F_n(x)$ — эмпирическая функция распределения, построенная по выборке $\vec{\xi}_n$. Из теоремы Гливенко (Теорема 11.1.1) мы знаем, что эта случайная величина с вероятностью 1 сходится к 0. Оказывается, при определенной нормировке величины $\sup\limits_{x\in\mathbb{R}} |F_n(x) - F_\xi(x)|$ ее пределом становится случайная величина, отличная от нуля, для которой удается описать ее распределение вероятностей. В результате этого возникает статистика, которую можно использовать для построения критерия согласия. Это построение опирается на следующую теорему, которую мы приводим без доказательства.

Теорема 12.2.1. — Колмогоров. Для любого $z \in \mathbb{R}$ имеет место сходимость

$$P\left(\sqrt{n}\sup_{x\in\mathbb{R}}|F_n(x) - F_\xi(x)| < z\right) \xrightarrow[n\to\infty]{} \mathcal{K}(z)\,,$$

где $\mathcal{K}(z)$ — ***функция Колмогорова***. Она определяется равенством

$$\mathcal{K}(z) = \begin{cases} \sum\limits_{k=-\infty}^{\infty}(-1)^k e^{-2k^2z^2}\,, & z > 0\,, \\ 0\,, & z \le 0\,. \end{cases}$$

R Функция Колмогорова является функцией распределения вероятностей, которое называется ***распределением Колмогорова*** и обозначается символом $\mathcal{K}$.

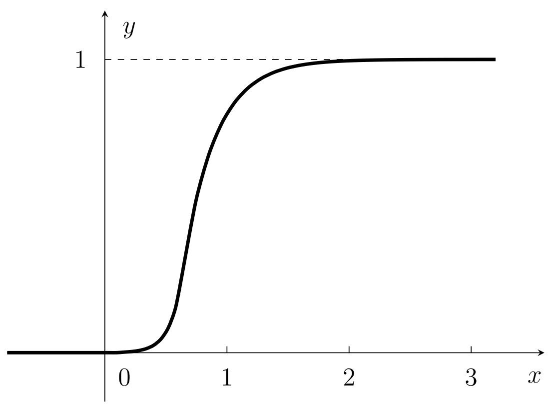


Рис. 12.1. График функции распределения Колмогорова $y = \mathcal{K}(x)$

Определение 12.2.1. Случайная величина, определенная равенством

$$D_n\big(\vec{\xi_n}\big) = \sqrt{n}\sup_{x\in\mathbb{R}}|F_n(x) - F_\xi(x)|\,,$$

называется ***статистикой Колмогорова***.

Теорема 12.2.1 утверждает, что последовательность $\left\{D_n\big(\vec{\xi}_n\big)\right\}$ сходится по распределению к случайной величине, распределенной по закону Колмогорова:

$$D_n\big(\vec{\xi}_n\big) \xrightarrow[n\to\infty]{d} \eta \sim \mathcal{K}\,.$$

Это означает, что при достаточно больших n можно пользоваться приближенным равенством

$$P\left(D_n\big(\vec{\xi}_n\big) < z\right) \approx \mathcal{K}(z)\,.$$

В результате приходим к следующему определению.

Определение 12.2.2. Критерием согласия Колмогорова с уровнем значимости α называется следующая функция от выборки:

$$\eta_n\big(\vec{\xi}_n\big) = \mathbf{1}_{\left\{D_n(\vec{\xi}_n) < z_{1-\alpha}\right\}}\,,$$

где $z_{1-\alpha} = \mathcal{K}^{-1}(1-\alpha)$ — квантиль порядка $1-\alpha$ распределения Колмогорова.

То, что уровень значимости критерия Колмогорова действительно равен α, следует из равенств:

$$\begin{aligned} P\left(\eta_n\big(\vec{\xi}_n\big) = 0 \,\middle|\, H_0\right) &= P\left(D_n\big(\vec{\xi}_n\big) \geq z_{1-\alpha} \,\middle|\, H_0\right) = \\ &= 1 - P\left(D_n\big(\vec{\xi}_n\big) < z_{1-\alpha} \,\middle|\, H_0\right) = \\ &= 1 - \mathcal{K}(z_{1-\alpha}) = 1 - (1-\alpha) = \alpha\,. \end{aligned}$$

Проверим состоятельность критерия Колмогорова. Для этого заметим, что из утверждения теоремы Гливенко следует сходимость

$$\sup_{x\in\mathbb{R}} |F_n(x) - F_0(x)| \xrightarrow[n\to\infty]{} \sup_{x\in\mathbb{R}} |F_\xi(x) - F_0(x)|\,.$$

При этом, если верна гипотеза H_1, то $\sup\limits_{x\in\mathbb{R}} |F_\xi(x) - F_0(x)| > 0$, так как $F_\xi(x) \not\equiv F_0(x)$. Следовательно,

$$D_n\big(\vec{\xi}_n\big) \xrightarrow[n\to\infty]{} \infty\,. \tag{12.1}$$

Пусть $B_n = \{D_n(\vec{\xi}_n) < z_{\alpha-1}\} \cdot H_1$. Тогда, в силу сходимости (12.1), $\prod\limits_{n=1}^{\infty} B_n = \varnothing$, а поскольку $B_n \supseteq B_{n+1}$ для любого $n \in \mathbb{N}$, по свойству непрерывности вероятности (Р-9) получаем

$$\lim_{n\to\infty} P(B_n) = P\left(\prod_{n=1}^{\infty} B_n\right) = P(\varnothing) = 0\,.$$

В результате

$$\beta = P\left(\eta_n(\vec{\xi}_n) = 1 \middle| H_1\right) = P\left(D_n(\vec{\xi}_n) < z_{1-\alpha} \middle| H_1\right) = \frac{P(B_n)}{P(H_1)} \xrightarrow[n\to\infty]{} 0\,.$$

На рис. 12.2 изображена схема работы критерия Колмогорова. Для того чтобы он принял значение 1, статистика Колмогорова

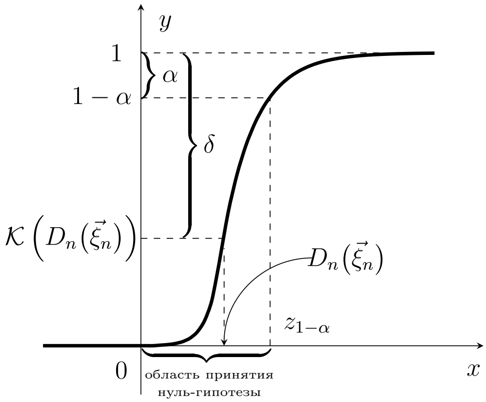


Рис. 12.2. Схема критерия Колмогорова

$D_n(\vec{\xi}_n)$ должна попасть в промежуток $(0; z_{1-\alpha})$ — область принятия нуль-гипотезы. При этом качество нуль-гипотезы тем выше, чем дальше от $z_{1-\alpha}$ — правой границы этой области — оказывается $D_n(\vec{\xi}_n)$. Величина $\delta = 1 - \mathcal{K}\left(D_n(\vec{\xi}_n)\right)$ называется ***критическим уровнем значимости*** критерия Колмогорова, соответствующим данному значению статистики $D_n(\vec{\xi}_n)$. Если положить $\alpha = \delta$, будет выполнено равенство $D_n(\vec{\xi}_n) = z_{1-\alpha}$, в результате чего критерий Колмогорова примет значение 0 и нуль-гипотезу нужно будет отвергнуть. Отсюда следует, что ***критический уровень значимости равен вероятности совершить ошибку, отвергнув***

нуль-гипотезу при данном значении статистики Колмогорова.

12.3. Критерий согласия Пирсона

Пусть $\{B_k\}_{k=1}^r$ — набор непересекающихся промежутков в $\mathbb{R}$. Обозначим $p_k^0 = P\big(\xi \in B_k \mid H_0\big)$. Будем предполагать, что $\left\{\xi \in \bigsqcup_{k=1}^r B_k\right\} = \Omega$ и $p_k^0 \neq 0\,, k = 1, \ldots, r$. Тогда $\sum_{k=1}^r p_k^0 = 1$.

Пусть $p_k = P\big(\xi \in B_k\big)$. Гипотеза H_0 имеет место тогда и только тогда, когда $p_k = p_k^0$ для всех $k = 1, \ldots, r$, а альтернативная гипотеза H_1 — тогда и только тогда, когда $p_k \neq p_k^0$ хотя бы при одном значении k от 0 до r.

Пусть $\nu_k = \sum\limits_{i=1}^n \mathbf{1}_{B_k}(\xi_i)$. Случайная величина ν_k дает количество элементов выборки $\vec{\xi}_n$, принадлежащих B_k. Рассмотрим случайную величину $\dfrac{\nu_k}{n}$ — частоту попадания компонент выборки $\vec{\xi}_n$ в промежуток B_k. Поскольку $E\nu_k = np_k$, имеем $E\dfrac{\nu_k}{n} = p_k$, а по слабому закону больших чисел $\dfrac{\nu_k}{n} \xrightarrow[n\to\infty]{P} p_k$. Таким образом, статистика $\dfrac{\nu_k}{n}$ представляет собой состоятельную и несмещенную оценку вероятности p_k. Критерий Пирсона основан на сравнении случайного вектора наблюдаемых частот $\dfrac{\vec{\nu}}{n} = \left(\dfrac{\nu_1}{n}, \ldots, \dfrac{\nu_r}{n}\right)^T$ и вектора вероятностей $(p_1^0, \ldots, p_r^0)^T$, соответствующих нуль-гипотезе.

Заметим, что для последовательности $\{\mathbf{1}_{B_k}(\xi_i)\}_{i=1}^\infty$ справедлива центральная предельная теорема (Теорема 10.6.1). Для дисперсии ν_k справедливо равенство

$$D\nu_k = D\sum_{i=1}^n \mathbf{1}_{B_k}(\xi_i) = nD\mathbf{1}_{B_k}(\xi) = n\big(E\mathbf{1}_{B_k}^2(\xi) - (E\mathbf{1}_{B_k}(\xi))^2\big) =$$

$$= n(p_k - p_k^2) = np_k(1 - p_k)\,,$$

поэтому по центральной предельной теореме имеет место сходимость

$$\frac{\nu_k - np_k}{\sqrt{np_k(1-p_k)}} \xrightarrow[n\to\infty]{d} N(0,1)\,.$$

Таким образом, распределение вероятностей случайного вектора

$$\vec{\eta}_n = \left(\frac{\nu_1 - np_1}{\sqrt{np_1(1-p_1)}}, \ldots, \frac{\nu_r - np_r}{\sqrt{np_r(1-p_r)}}\right)$$

при неограниченном увеличении объема выборки $\vec{\xi}_n$ стремится к r-мерному гауссовскому (см. Приложение C, где рассматриваются свойства гауссовских случайных векторов). Хотя распределения вероятностей компонент вектора $\vec{\eta}_n$ стремятся к стандартному гауссовскому, предельное распределение вероятностей $\vec{\eta}_n$ не является сферически-симметрическим. Более того, оно оказывается вырожденным. Это объясняется тем, что случайные величины ν_k зависимы, поскольку связаны соотношением

$$\sum_{k=1}^{r} \nu_k = n\,.$$

Оказывается, если компоненты вектора $\vec{\eta}_n$ подходящим образом нормировать, а именно перейти к вектору

$$\vec{\zeta}_n = \left(\frac{\nu_1 - np_1}{\sqrt{np_1}}, \ldots, \frac{\nu_r - np_r}{\sqrt{np_r}}\right)\,,$$

предельное распределение вероятностей получающегося r-мерного вектора сосредоточено на некотором $r-1$-мерном подпространстве пространства $\mathbb{R}^r$ и является на нем сферически-симметрическим. Этот факт и составляет содержание доказательства теоремы Пирсона, которая лежит в основе построения критерия Пирсона.

Определение 12.3.1. Пусть $\vec{\theta} = (\theta_1, \ldots, \theta_n)^T \sim N(\overline{0}, E)$. Распределение вероятностей случайной величины $\xi = |\theta|^2 = \sum\limits_{i=1}^{n} \theta_i^2$ называется распределением ***хи-квадрат с n степенями свободы***. При этом пишут $\xi \sim \chi_n^2$.

R На самом деле распределение хи-квадрат с n степенями свободы является частным случаем гамма-распределения: $\chi^2_n \sim \Gamma\left(\frac{n}{2}, 2\right)$ (см. Приложение D, где изложены базовые свойства распределения хи-квадрат и доказательство следующей теоремы).

Теорема 12.3.1. — Пирсон. При $n \to \infty$ распределение вероятностей суммы $\sum\limits_{k=1}^{r} \dfrac{(\nu_k - np_k)^2}{np_k}$ сходится к распределению χ^2_{r-1}.

Определение 12.3.2. ***Статистика Пирсона*** определяется формулой

$$\gamma_n\big(\vec{\xi}_n\big) = \sum_{k=1}^{r} \frac{(\nu_k - np_k^0)^2}{np_k^0}\,.$$

Критерий Пирсона с уровнем значимости $\boldsymbol{\alpha}$ определяется равенством

$$\eta\big(\vec{\xi}_n\big) = \mathbf{1}_{\{\gamma_n(\vec{\xi}_n) < z_{1-\alpha}\}}\,,$$

где $z_{1-\alpha}$ — квантиль порядка $1-\alpha$ распределения χ^2_{r-1}.

Убедимся в том, что уровень значимости критерия Пирсона действительно равен α. Для этого заметим, что если гипотеза H_0 верна, то $p_k^0 = p_k$ для всех k, а значит для статистики Пирсона выполняется равенство

$$\gamma_n\big(\vec{\xi}_n\big) = \sum_{k=1}^{r} \frac{(\nu_k - np_k)^2}{np_k}$$

и при достаточно большом объеме выборки, по теореме Пирсона, можно считать, что она распределена по закону χ^2_{r-1}. В результате получаем

$$\alpha = P\left(\gamma_n\big(\vec{\xi}_n\big) \geq z_{1-\alpha} \,\middle|\, H_0\right) = 1 - F_{\chi^2_{r-1}}(z_{1-\alpha}) = 1 - (1-\alpha) = \alpha\,.$$

Проверим состоятельность критерия Пирсона. Если верна гипотеза H_1, имеем:

$$\sum_{k=1}^{r} \frac{(\frac{\nu_k}{n} - p_k^0)^2}{p_k^0} \xrightarrow[n\to\infty]{} \sum_{k=1}^{r} \frac{(p_k - p_k^0)^2}{p_k^0} > 0\,.$$

Следовательно,

$$\gamma_n\left(\vec{\xi}_n\right) = \sum_{k=1}^{r} \frac{(\nu_k - np_k^0)^2}{np_k^0} = n \cdot \sum_{k=1}^{r} \frac{(\frac{\nu_k}{n} - p_k^0)^2}{p_k^0} \xrightarrow[n\to\infty]{} \infty\,,$$

причем, начиная с некоторого номера n, выполняется неравенство

$$\gamma_n\left(\vec{\xi}_n\right) < \gamma_{n+1}\left(\vec{\xi}_{n+1}\right). \tag{12.2}$$

Далее,

$$\beta = P\left(\gamma_n\left(\vec{\xi}_n\right) < z_{1-\alpha} \,\middle|\, H_1\right) = \frac{P(A_n)}{P(H_1)}\,,$$

где $A_n = \left\{\gamma_n\left(\vec{\xi}_n\right) < z_{1-\alpha}\right\} \cdot H_1$. Из неравенства (12.2) следуют вложения $A_n \supseteq A_{n+1}$, а значит, по свойству непрерывности вероятности (Р-9), получаем

$$\lim_{n\to\infty} P(A_n) = P\left(\prod_{n=1}^{\infty} A_n\right) = P(\varnothing) = 0\,.$$

Следовательно, $\beta \xrightarrow[n\to\infty]{} 0$.

В приложении Е содержатся задания лабораторной работы «Исследование распределения вероятностей по реализации выборки. Критерии согласия» (раздел Е.2), которую предлагается выполнить в вычислительной среде *MatLab*.

Библиография

1. Бородин, А. Н. Элементарный курс теории вероятностей и математической статистики / А. Н. Бородин. – 3-е изд., стер. – СПб. : Лань, 2004. – 224 с. – ISBN 978-5-8114-0442-1.

2. Вентцель, Е. С. Теория вероятностей : учебник / Е. С. Вентцель. – 11-е изд., стер. – М. : Кнорус, 2018. – 664 с. – ISBN 978-5-4060-0476-0.

3. Гмурман, В. Е. Руководство к решению задач по теории вероятностей и математической статистике : учебное пособие для студентов вузов / В. Е. Гмурман. – 9-е изд., стер. – М. : Высш. шк., 2004. – 404 с. – ISBN 5-06-004212-X.

4. Гнеденко, Б. В. Курс теории вероятностей / Б. В. Гнеденко. – 8-е изд., испр. и доп. – М. : Едиториал УРСС, 2005. – 448 с. – ISBN 5-354-01091-8.

5. Севастьянов, Б. А. Сборник задач по теории вероятностей : учебное пособие для вузов / Б. А. Севастьянов, В. П. Чистяков, А. М. Зубков. – СПб. : Лань, 2009. – 320 с. – ISBN 978-5-8114-0975-4.

6. Колмогоров, А. Н. Элементы теории функций и функционального анализа / А. Н. Колмогоров, С. В. Фомин. – 7-е изд., стер. – М. : Изд-во Моск. ун-та, 2019. – 572 с. – ISBN 978-5-9221-0266-7.

7. Свешников, А. А. Сборник задач по теории вероятностей, математической статистике и теории случайных функций / под общ. ред. А. А. Свешникова. – 5-е изд., стер. – СПб. : Лань, 2013. – 448 с. – ISBN 978-5-8114-0708-8.

8. Чистяков, В. П. Курс теории вероятностей / В. П. Чистяков. – 5-е изд., испр. – М. : Дрофа, 2007. – 256 с. – ISBN 978-5-358-03022-0.

9. Ширяев, А. Н. Вероятность : в 2-х кн. / А. Н. Ширяев — 4-е изд., перераб. и доп. — М. : МЦНМО, 2007. – Кн. 1. — 552 с. — ISBN 978-5-94057-105-6. – Кн. 2. — 416 с. — ISBN 978-5-94057-106-3.

Приложение А. Элементы комбинаторики

При вычислении вероятностей в схеме с конечным числом равновероятных исходов для определения количества элементов того или иного множества используют теоремы и формулы комбинаторики — раздела математики, изучающего методы подсчета комбинаций различных объектов. Рассмотрим некоторые из них.

Базовыми в комбинаторике являются две элементарные теоремы — «Правило суммы» и «Правило произведения».

Правило суммы

Если объект a можно выбрать n способами, а объект b — m способами, то a или b можно выбрать $n+m$ способами.

■ Пример A.0.1. Если имеется n яблок и m груш, то яблоко или грушу (т. е. один фрукт из этой совокупности) можно выбрать $n+m$ способами. ■

Правило произведения

Если объект a можно выбрать n способами и при каждом выборе a объект b можно выбрать m способами, то упорядоченную пару (a, b) можно выбрать $n \cdot m$ способами.

■ Пример A.0.2. Если имеется n яблок и m груш, то яблоко и грушу (т. е. два фрукта) можно выбрать $n \cdot m$ способами. ■

Определение А.0.1. ***Размещением из n элементов по m*** называется упорядоченная m-ка, состоящая из попарно различных элементов некоторого n-элементного множества. Число всевозможных размещений из n по m обозначают символом A_n^m.

■ Пример А.0.3. Рассмотрим множество $A = \{\square, \bigcirc, \triangle\}$. Выпишем все размещения из его трех элементов по два:

$$(\square, \bigcirc), (\bigcirc, \square), (\square, \triangle), (\triangle, \square), (\bigcirc, \triangle), (\triangle, \bigcirc)\,.$$

Заметим, что пары, отличающиеся друг от друга порядком элементов, представляют собой разные размещения. ■

С помощью правила произведения доказывается формула

$$\boxed{A_n^m = n(n-1)\cdots(n-m+1)\,. \qquad \text{(A.1)}}$$

Доказательство. Пусть имеется n-элементное множество и нужно построить какое-нибудь размещение $(x_1, x_2, \ldots, x_m)$ из этих n элементов по m, где $0 < m \leq n$. Первый элемент размещения можно выбрать n способами. Если x_1 выбран, второй элемент можно выбрать из оставшихся элементов $n-1$ способами, следовательно, упорядоченную пару (x_1, x_2), по правилу произведения, можно выбрать $n(n-1)$ способами. Если выбрана пара (x_1, x_2), то существует $n-2$ способа выбрать следующий элемент размещения x_3 из оставшихся элементов исходного множества. Следовательно, по правилу произведения тройку (x_1, x_2, x_3) можно выбрать $n(n-1)(n-2)$ способами. Продолжая применять правило произведения индуктивно, получаем равенство (А.1). ■

Определение А.0.2. ***Перестановкой из n элементов*** называется упорядоченная n-ка, состоящая из попарно различных элементов некоторого n-элементного множества. Число всевозможных перестановок из n элементов обозначают P_n.

Перестановка из n элементов является размещением из n по n, поэтому из формулы (А.1) следует равенство

$$P_n = n! \tag{A.2}$$

Определение A.0.3. ***Сочетанием из n элементов по m*** называется произвольное m-элементное подмножество, состоящее из попарно различных элементов некоторого n-элементного множества. Число всевозможных сочетаний из n по m обозначают символом C_n^m.

■ Пример A.0.4. Для множества $A = \{\square, \bigcirc, \triangle\}$, рассмотренного в примере A.0.3, выпишем все сочетания из его трех элементов по два:

$$\{\square, \bigcirc\}, \{\square, \triangle\}, \{\bigcirc, \triangle\}\,.$$

Заметим, что $\{\square, \bigcirc\}$ и $\{\bigcirc, \square\}$ — это одно и то же сочетание. ■

Для числа сочетаний из n по m справедливо равенство

$$C_n^m = \frac{A_n^m}{P_m} = \frac{n(n-1)\cdots(n-m+1)}{m!} = \frac{n!}{m!(n-m)!}\,. \tag{A.3}$$

Доказательство. Из элементов одного сочетания из n по m можно построить P_m разных перестановок, т. е. размещений из n по m. При этом из разных сочетаний нельзя построить одинаковые размещения. Отсюда следует равенство $C_n^m P_m = A_n^m$, а значит равенство (A.3). ■

R Кроме размещений, перестановок и сочетаний, состоящих из попарно различных элементов, рассматривают ***размещения, перестановки и сочетания с повторениями***, где один и тот же элемент может участвовать несколько раз. Для обозначения числа размещений с повторениями из n элементов по m часто используют символ $\overline{A}_n^m$, при этом из правила произведения следует

$$\overline{A}_n^m = n^m\,. \tag{A.4}$$

Под перестановкой с повторениями понимается упорядоченный набор из n элементов, каждый из которых относится к одному из m типов. Элементы одного типа неразличимы, так что перестановки, отличающиеся друг от друга положением однотипных элементов, отождествляются. Пусть k_i — количество элементов i-го типа, так что $k_1 + k_2 + \cdots + k_m = n$, тогда для $\overline{P}_n$ — числа перестановок с повторениями из n элементов — справедливо равенство

$$\boxed{\overline{P}_n = \frac{n!}{k_1!k_2!\ldots k_m!}\,. \qquad \text{(A.5)}}$$

Число сочетаний с повторениями из n элементов по m обозначают $\overline{C}_n^m$. Для него можно доказать формулу

$$\boxed{\overline{C}_n^m = \frac{(n+m-1)!}{m!(n-1)!}\,. \qquad \text{(A.6)}}$$

Приложение B. Гамма-функция Эйлера

Определение B.0.1. ***Гамма-функция Эйлера*** определяется равенством

$$\Gamma(x) = \int\limits_0^{+\infty} t^{x-1}e^{-t}\,dt\,. \tag{B.1}$$

Интеграл (B.1) представляет собой сходящийся несобственный интеграл 1-го рода при $x \geq 1$. При $x < 1$ подинтегральная функция имеет особенность в точке $t = 0$ и интеграл (B.1) представляет собой несобственный интеграл смешанного типа:

$$\int\limits_0^{+\infty} t^{x-1}e^{-t}\,dt = \int\limits_0^{a} t^{x-1}e^{-t}\,dt + \int\limits_a^{+\infty} t^{x-1}e^{-t}\,dt\,, \quad a > 0\,,$$

где первое слагаемое — несобственный интеграл 2-го рода, сходящийся при $0 < x < 1$. Поскольку второе слагаемое — сходящийся несобственный интеграл 1-го рода при любом $x \in \mathbf{R}$, гамма-функция определена при $x \in (0; +\infty)$.

Можно доказать, что гамма-функция бесконечно дифференцируема на своей области определения. Одним из ее важных свойств является следующее:

$$\Gamma(x+1) = x\Gamma(x)\,, \quad x > 0\,. \tag{B.2}$$

Действительно, для любого $x > 0$, интегрируя по частям, получаем:

$$\Gamma(x+1) = \int\limits_0^{+\infty} t^x e^{-t}\,dt = -\int\limits_0^{+\infty} t^x\,de^{-t} = -t^x e^{-t}\Big|_0^{+\infty} + \int\limits_0^{+\infty} e^{-t}\,dt^x = $$
$$= x\int\limits_0^{+\infty} t^{x-1}e^{-t}\,dt = x\Gamma(x)\,.$$

Формулу (В.2) называют ***формулой понижения*** для гамма-функции.

Поскольку при $x = 1$ имеем

$$\Gamma(1) = \int\limits_0^{+\infty} e^{-t}\,dt = -e^{-t}\Big|_0^{+\infty} = 1\,,$$

из формулы понижения следует, что

$$\Gamma(n+1) = n\Gamma(n) = n(n-1)\Gamma(n-1) = \cdots = n!\,,\quad n \in \mathbb{N}\,.$$

Таким образом, гамма-функция представляет собой бесконечно дифференцируемое продолжение факториала на множество всех положительных действительных чисел.

Отметим, что при $x = \dfrac{1}{2}$, используя интеграл Пуассона, получаем

$$\Gamma\left(\frac{1}{2}\right) = \int\limits_0^{+\infty} \frac{e^{-t}}{\sqrt{t}}\,dt = 2\int\limits_0^{+\infty} e^{-t}\,d\sqrt{t} = 2\int\limits_0^{+\infty} e^{-s^2}\,ds = \int\limits_{-\infty}^{+\infty} e^{-s^2}\,ds = \sqrt{\pi}\,.$$

Отсюда, используя формулу понижения, получаем еще одну формулу:

$$\Gamma\left(\frac{2n+1}{2}\right) = \left(\frac{2n-1}{2}\right)\left(\frac{2n-3}{2}\right)\ldots\frac{1}{2}\sqrt{\pi} = \frac{(2n-1)!!}{2^n}\sqrt{\pi}\,.$$

Приложение C. Многомерное нормальное распределение

C.1. Невырожденное многомерное нормальное распределение

Определение C.1.1. Случайный вектор $\vec{\eta} = (\eta_1, \eta_2, \ldots, \eta_n)$ называется ***сферически-симметрическим гауссовским*** или ***сферически-симметрическим нормальным***, если случайные величины η_i независимы и $\eta_i \sim N(0,1)\,, i = 1, 2, \ldots, n$. При этом пишут $\vec{\eta} \sim N(\overline{0}, E)$.

По теореме 6.5.3 из этого определения следует, что плотность распределения сферически-симметрического нормального случайного вектора $\vec{\eta}$ или плотность совместного распределения его компонент — случайных величин $\eta_1, \eta_2, \ldots, \eta_n$, — имеет вид

$$f_{\vec{\eta}}(\overline{x}) = f_{\eta_1,\eta_2,\ldots,\eta_n}(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} f_{\eta_i}(x_i) = \\ = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_i^2}{2}} = \frac{1}{(2\pi)^{n/2}} \exp\left[-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right].$$

Таким образом,

$$f_{\vec{\eta}}(\overline{x}) = \frac{1}{(2\pi)^{n/2}} \exp\left[-\frac{1}{2}(\overline{x}, \overline{x})\right]. \tag{C.1}$$

Это распределение вероятностей называется ***сферически-симметрическим***, потому что поверхности уровня его плотности представляют собой n-мерные сферы (если $n = 2$, то это окружности, если $n = 3$, то это обычные трехмерные сферы).

Предложение C.1.1. Пусть $\vec{\eta} \sim N(\overline{0}, E)$ — n-мерный случайный вектор, $\vec{\xi} = B\vec{\eta} + \overline{a}$, где $\overline{a} \in \mathbb{R}^n$, B — невырожденная матрица $n \times n$. Тогда

$$f_{\vec{\xi}}(\overline{x}) = \frac{1}{(2\pi)^{n/2}\sqrt{|\Lambda|}} \exp\left[-\frac{1}{2}\left(\Lambda^{-1}(\overline{x} - \overline{a}), \overline{x} - \overline{a}\right)\right], \tag{C.2}$$

где $\Lambda = BB^T$.

Доказательство. Линейное преобразование $\overline{x} = \overline{g}(\overline{y}) = B\overline{y} + \overline{a}$ удовлетворяет условиям предложения 6.3.2, поэтому, по формуле (6.8), получаем

$$\begin{aligned} f_{\vec{\xi}}(\overline{x}) &= \frac{1}{|J(\overline{g})|} f_{\vec{\eta}}\big(g^{-1}(\overline{x})\big) = \\ &= \frac{1}{(2\pi)^{n/2}|B|} \exp\left[-\frac{1}{2}\left(B^{-1}(\overline{x} - \overline{a}), B^{-1}(\overline{x} - \overline{a})\right)\right] = \\ &= \frac{1}{(2\pi)^{n/2}|B|} \exp\left[-\frac{1}{2}\left((B^{-1})^T B^{-1}(\overline{x} - \overline{a}), \overline{x} - \overline{a}\right)\right]. \end{aligned}$$

Матрица $\Lambda = BB^T$ является положительно определенной, так как для любого $\overline{x} \in \mathbb{R}^n$

$$(\Lambda\overline{x}, \overline{x}) = (BB^T\overline{x}, \overline{x}) = (B^T\overline{x}, B^T\overline{x}) \geq 0,$$

при этом $(\Lambda\overline{x}, \overline{x}) = 0 \ \Leftrightarrow\ B^T\overline{x} = \overline{0} \ \Leftrightarrow\ \overline{x} = \overline{0}$ в силу невырожденности B. Следовательно, матрица Λ невырождена, при этом

$$\Lambda^{-1} = (BB^T)^{-1} = (B^{-1})^T B^{-1}, \quad |\Lambda| = |BB^T| = |B| \cdot |B^T| = |B|^2.$$

В результате получаем формулу (C.2). ■

Определение C.1.2. Случайный вектор $\vec{\xi}$ называется (невырожденным) ***гауссовским*** или ***нормальным с параметрами*** $\boldsymbol{\overline{a}}$, $\boldsymbol{\Lambda}$, где $\overline{a} \in \mathbb{R}^n$, а Λ — положительно определенная матрица $n \times n$, если плотность его распределения вероятностей имеет вид (C.2).

При этом пишут $\vec{\xi} \sim N(\overline{a}, \Lambda)$.

Ⓡ Сферически-симметрический нормальный вектор представляет собой частный случай нормального случайного вектора, где $\overline{a} = \overline{0}$, $\Lambda = E$ — единичная матрица.

Из предложения С.1.1 вытекает следующее

Следствие С.1.2. Пусть $\overline{a} \in \mathbb{R}^n$, Λ — положительно-определенная матрица $n \times n$. Тогда $\vec{\xi} \sim N(\overline{a}, \Lambda) \;\Leftrightarrow\; \vec{\xi} = B\vec{\eta} + \overline{a}$, где $\vec{\eta} \sim N(\overline{0}, E)$, B — невырожденная матрица, такая, что $\Lambda = BB^T$.

Доказательство. Достаточность утверждает предложение С.1.1. Докажем необходимость. Пусть Λ — положительно определенная матрица. Известно, что существует ортогональное преобразование T, приводящее ее к диагональному виду:

$$\Lambda = T\Theta T^T\,, \quad \Theta = \mathrm{diag}(\theta_1, \theta_2, \ldots, \theta_n)\,,$$

где $\mathrm{diag}(\theta_1, \theta_2, \ldots, \theta_n)$ — матрица $n \times n$, все элементы которой, кроме диагональных, равных $\theta_1, \theta_2, \ldots, \theta_n$, равны нулю, причем $\theta_i > 0$ для $i = 1, 2, \ldots, n$. Положим $B = T\sqrt{\Theta}$, где $\sqrt{\Theta} = \mathrm{diag}(\sqrt{\theta_1}, \sqrt{\theta_2}, \ldots, \sqrt{\theta_n})$. Тогда B — невырожденная матрица, связанная с Λ равенством $\Lambda = BB^T$. Определим случайный вектор $\vec{\eta}$ равенством

$$\vec{\eta} = \overline{g}(\vec{\xi}) = B^{-1}(\vec{\xi} - \overline{a}) = B^{-1}\vec{\xi} - B^{-1}\overline{a}\,.$$

Тогда, в силу предложения 6.3.2, по формуле (6.8), пользуясь тем, что $\overline{g}^{-1}(\overline{x}) = B\overline{x} + \overline{a}$, $J(\overline{g}) = |B^{-1}| = \dfrac{1}{|B|} = \dfrac{1}{\sqrt{|\Lambda|}}$, $\Lambda^{-1} = (B^T)^{-1}B^{-1}$, получаем

$$f_{\vec{\eta}}(\overline{x}) = \frac{1}{|J(\overline{g})|} f_{\vec{\xi}}\left(\overline{g}^{-1}(\overline{x})\right) = \frac{1}{(2\pi)^{n/2}} \exp\left[-\frac{1}{2}\left(\Lambda^{-1}B\overline{x}, B\overline{x}\right)\right] = \\ = \frac{1}{(2\pi)^{n/2}} \exp\left[-\frac{1}{2}\left(\overline{x}, \overline{x}\right)\right]\,,$$

т. е. $\eta \sim N(\overline{0}, E)$, при этом $\vec{\xi} = B\vec{\eta} + \overline{a}$. ■

Следствие С.1.2 позволяет выяснить смысл параметров многомерного нормального распределения вероятностей.

Следствие C.1.3. Если $\vec{\xi} = (\xi_1, \xi_2, \ldots, \xi_n)^T \sim N(\overline{a}, \Lambda)$, то $\overline{a} = (a_1, a_2, \ldots, a_n)^T = E\vec{\xi}$, а Λ — ковариационная матрица вектора $\vec{\xi}$, при этом для любого $i = 1, 2, \ldots, n$ $\xi_i \sim N(a_i, \sigma_i^2)$, где $\sigma_i^2 = \lambda_{ii}$.

Доказательство. Пусть $\vec{\xi} \sim N(\overline{a}, \Lambda)$. Тогда $\vec{\xi} = B\vec{\eta} + \overline{a}$, где $\vec{\eta} \sim N(\overline{0}, E)$, а $BB^T = \Lambda$. В результате для любого i справедливо равенство $\xi_i = \sum\limits_{k=1}^{n} b_{ik}\eta_k + a_i$. То, что линейная функция от гауссовских случайных величин является гауссовской случайной величиной? следует из предложения 5.5.3 и примеров 7.1.7, 7.2.2 и 9.2.3. При этом для моментов ξ_i справедливы равенства

$$E\xi_i = \sum_{k=1}^{n} b_{ik} E\eta_k + a_i = a_i\,, \quad D\xi_i = \sigma^2 = \sum_{k=1}^{n} b_{ik}^2 = \lambda_{ii}\,.$$

Следовательно, $\xi_i \sim N(a_i, \sigma_i^2)$.

Таким образом, $\overline{a} = E\vec{\xi}$. Используя свойства ковариационного момента (C-I), (C-II), (C-IV) и (C-VI), получаем:

$$\begin{aligned}\mathrm{Cov}\,(\xi_i, \xi_j) &= \mathrm{Cov}\,\left(\sum_{k=1}^{n} b_{ik}\eta_k + a_i, \sum_{m=1}^{n} b_{jm}\eta_m + a_j\right) = \\ &= \sum_{k=1}^{n}\sum_{m=1}^{n} b_{ik}b_{jm}\mathrm{Cov}\,(\eta_k, \eta_m) = \sum_{k=1}^{n} b_{ik}b_{jk} = \lambda_{ij}\,.\end{aligned}$$

■

Следствие C.1.4. Гауссовские случайные величины независимы тогда и только тогда, когда некоррелированы.

R Независимость случайных величин влечет за собой некоррелированность, при этом обратное в общем случае неверно (см. замечание на стр. 120).

Доказательство. Достаточно доказать, что для гауссовских случайных величин из некоррелированности следует независимость. Пусть

$\xi_i \sim N(a_i, \sigma_i^2)\,, i = 1, 2, \ldots, n$ и случайные величины ξ_i некоррелированы. Тогда $\vec{\xi} = (\xi_1, \xi_2, \ldots, \xi_n)^T \sim N(\overline{a}, \Lambda)$, где $\overline{a} = (a_1, a_2, \ldots, a_n)^T$, а ковариационная матрица является диагональной:

$$\Lambda = \operatorname{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2)\,.$$

Тогда

$$\Lambda^{-1} = \operatorname{diag}\left(\frac{1}{\sigma_1^2}, \frac{1}{\sigma_2^2}, \ldots, \frac{1}{\sigma_n^2}\right)\,.$$

Следовательно, плотность совместного распределения случайных величин $\xi_1, \xi_2, \ldots, \xi_n$ имеет вид:

$$\begin{aligned} f_{\vec{\xi}}(\overline{x}) &= \frac{1}{(2\pi)^{n/2}\sqrt{|\Lambda|}} \exp\left[-\frac{1}{2}\left(\Lambda^{-1}(\overline{x}-\overline{a}), \overline{x}-\overline{a}\right)\right] = \\ &= \frac{1}{(2\pi)^{n/2}\prod_{i=1}^{n}\sigma_i} \exp\left[-\frac{1}{2}\sum_{i=1}^{n}\frac{(x_i-a_i)^2}{\sigma_i^2}\right] = \\ &= \prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma_i}\exp\left[-\frac{(x_i-a_i)^2}{2\sigma_i^2}\right]\,. \end{aligned} \qquad (C.3)$$

Пользуясь формулой (6.6), для любого $j = \overline{1, n}$ получаем

$$f_{\xi_j}(x) = \underset{\mathbb{R}^{n-1}}{\int \cdots \int} f_{\xi_1,\ldots,\xi_n}(x_1, \ldots, \underset{j}{x}, \ldots, x_n)\, dx_1 \cdots < dx_j > \ldots dx_n =$$

$$= \frac{1}{\sqrt{2\pi}\sigma_j}\exp\left[-\frac{(x_j-a_j)^2}{2\sigma_j^2}\right]\prod_{i\neq j}\int_{\mathbb{R}}\frac{1}{\sqrt{2\pi}\sigma_i}\exp\left[-\frac{(x_i-a_i)^2}{2\sigma_i^2}\right]dx_i =$$

$$= \frac{1}{\sqrt{2\pi}\sigma_j}\exp\left[-\frac{(x_j-a_j)^2}{2\sigma_j^2}\right]\,.$$

Здесь сначала из-под $n-1$ кратного интеграла выносится множитель, зависящий от x_j, а значит, не зависящий от переменных интегрирования. Далее при переходе к повторному интегралу $n-1$-кратный интеграл распадается в произведение интегралов, каждый из которых равен 1 т. к. под интегралом плотность одномерного нормального распределения. При этом равенство (C.3) означает, что плотность совместного распределения вероятностей случайных величин ξ_i распадается в произведение плотностей их маргинальных распределений:

$$f_{\vec{\xi}}(\overline{x}) = \prod_{i=1}^{n} f_{\xi_i}(x_i)\,.$$

По теореме 6.13 это эквивалентно их независимости. ■

C.2. Невырожденное двумерное нормальное распределение

Рассмотрим подробнее свойства двумерного нормального вектора. Пусть

$$\xi_1 \sim N(a_1, \sigma_1^2)\,, \quad \xi_2 \sim N(a_2, \sigma_2^2)\,, \quad \rho = \mathrm{Cov}\,(\xi_1, \xi_2)\,.$$

Тогда $\vec{\xi} = (\xi_1, \xi_2) \sim N(\overline{a}, \Lambda)$, где $\overline{a} = (a_1, a_2)$, а ковариационная матрица $\vec{\xi}$ имеет вид

$$\Lambda = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

Отсюда следуют равенства

$$|\Lambda| = \sigma_1^2\sigma_2^2(1-\rho^2)\,, \quad \Lambda^{-1} = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)} \begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix}.$$

В результате плотность совместного распределения ξ_1 и ξ_2 имеет вид:

$$\begin{aligned} f_{\vec{\xi}}(x_1, x_2) &= \frac{1}{(2\pi)^{n/2}\sqrt{|\Lambda|}} \exp\left[-\frac{1}{2}\left(\Lambda^{-1}(\overline{x}-\overline{a}), \overline{x}-\overline{a}\right)\right] = \\ &= \frac{1}{2\pi\sqrt{1-\rho^2}\sigma_1\sigma_2} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{(x_1-a_1)^2}{\sigma_1^2} - \right.\right. \\ &\qquad \left.\left. -\frac{2\rho(x_1-a_1)(x_2-a_2)}{\sigma_1\sigma_2} + \frac{(x_2-a_2)^2}{\sigma_2^2}\right)\right]. \end{aligned} \tag{C.4}$$

Линии уровня плотности совместного распределения ξ_1 и ξ_2 представляют собой эллипсы, заданные уравнениями вида

$$\frac{1}{2(1-\rho^2)}\left(\frac{(x_1-a_1)^2}{\sigma_1^2} - \frac{2\rho(x_1-a_1)(x_2-a_2)}{\sigma_1\sigma_2} + \frac{(x_2-a_2)^2}{\sigma_2^2}\right) = C\,.$$

Форма этих эллипсов определяется коэффициентом корреляции. Эллипс, получающийся при $C = 4$, называют эллипсом рассеивания случайного вектора $\vec{\xi}$. Обозначим через $\mathcal{E}$ область, ограниченную этим эллипсом, т. е. заданную неравенством

$$\frac{1}{2(1-\rho^2)}\left(\frac{(x_1-a_1)^2}{\sigma_1^2}-\frac{2\rho(x_1-a_1)(x_2-a_2)}{\sigma_1\sigma_2}+\frac{(x_2-a_2)^2}{\sigma_2^2}\right)\le 4$$
$$\Updownarrow$$
$$\frac{1}{2}\left(\Lambda^{-1}(\overline{x}-\overline{a}),\overline{x}-\overline{a}\right)\le 4\,.$$

Найдем вероятность события $\{\vec{\xi}\in\mathcal{E}\}$. Так же, как в доказательстве следствия С.1.2, представим ковариационную матрицу случайного вектора $\vec{\xi}$ в виде $\Lambda = BB^T$, где B невырожденная матрица $n\times n$. Получим:

$$P\big(\vec{\xi}\in\mathcal{E}\big)=\iint\limits_{\mathcal{E}} f_{\vec{\xi}}(x_1,x_2)\,dx_1dx_2=$$
$$=\frac{1}{2\pi\sqrt{|\Lambda|}}\iint\limits_{(\Lambda^{-1}(\overline{x}-\overline{a}),\overline{x}-\overline{a})\le 8}\exp\left[-\frac{1}{2}\left(\Lambda^{-1}(\overline{x}-\overline{a}),\overline{x}-\overline{a}\right)\right]d\overline{x}=$$
$$=\frac{1}{2\pi|B|}\iint\limits_{(B^{-1}(\overline{x}-\overline{a}),B^{-1}(\overline{x}-\overline{a}))\le 8}\exp\left[-\frac{1}{2}\left(B^{-1}(\overline{x}-\overline{a}),B^{-1}(\overline{x}-\overline{a})\right)\right]d\overline{x}\,.$$

Делая в интеграле замену $\overline{y}=B^{-1}(\overline{x}-\overline{a}) \ \Leftrightarrow\ \overline{x}=B\overline{y}+\overline{a}$, при которой $d\overline{x}=|B|d\overline{y}$, получаем:

$$P\big(\vec{\xi}\in\mathcal{E}\big)=\frac{1}{2\pi}\iint\limits_{(\overline{y},\overline{y})\le 8}e^{-\frac{1}{2}(\overline{y},\overline{y})}d\overline{y}=\frac{1}{2\pi}\iint\limits_{y_1^2+y_2^2\le 8}e^{-\frac{1}{2}(y_1^2+y_2^2)}dy_1dy_2\,.$$

Делая полярную замену в двойном интеграле, получаем:

$$P\big(\vec{\xi}\in\mathcal{E}\big)=\frac{1}{2\pi}\int\limits_0^{2\pi}d\varphi\int\limits_0^{2\sqrt{2}}e^{-\frac{\rho^2}{2}}\rho\,d\rho=-e^{-\frac{\rho^2}{2}}\Big|_0^{2\sqrt{2}}=1-e^{-4}\approx 0,98\,.$$

Таким образом, с вероятностью, близкой к единице, случайный вектор $\vec{\xi}$ попадает в область, ограниченную эллипсом рассеивания.

Построим кривую среднеквадратической регрессии одной компоненты двумерного гауссовского случайного вектора на другую. Для этого найдем плотность условного распределения $f_{\xi_1|\xi_2}(x|\xi_2 = y)$.

$$
\begin{aligned}
&f_{\xi_1|\xi_2}(x_1|\xi_2 = x_2) = \frac{f_{\xi_1\xi_2}(x_1, x_2)}{f_{\xi_2}(x_2)} = \\
&= \frac{\sqrt{2\pi}\sigma_2 \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{(x_1-a_1)^2}{\sigma_1^2} - \frac{2\rho(x_1-a_1)(x_2-a_2)}{\sigma_1\sigma_2} + \frac{(x_2-a_2)^2}{\sigma_2^2}\right)\right]}{2\pi\sqrt{1-\rho^2}\sigma_1\sigma_2 \exp\left[\frac{(x_2-a_2)^2}{2\sigma_2^2}\right]} = \\
&= \frac{1}{\sqrt{2\pi(1-\rho^2)}\sigma_1} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{(x_1-a_1)^2}{\sigma_1^2} - \right.\right. \\
&\qquad\qquad \left.\left. - \frac{2\rho(x_1-a_1)(x_2-a_2)}{\sigma_1\sigma_2} + \rho^2\frac{(x_2-a_2)^2}{\sigma_2^2}\right)\right] = \\
&= \frac{1}{\sqrt{2\pi(1-\rho^2)}\sigma_1} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{x_1-a_1}{\sigma_1} - \rho\frac{x_2-a_2}{\sigma_2}\right)^2\right] = \\
&= \frac{1}{\sqrt{2\pi(1-\rho^2)}\sigma_1} \exp\left[-\frac{1}{2(1-\rho^2)\sigma_1^2}\left(x_1 - a_1 - \rho\frac{\sigma_1}{\sigma_2}(x_2-a_2)\right)^2\right] = \\
&= \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x_1-a)^2}{2\sigma^2}\right],
\end{aligned}
$$

где $a = a_1 + \rho\dfrac{\sigma_1}{\sigma_2}(x_2 - a_2)$, $\sigma = \sigma_1\sqrt{1-\rho^2}$. Это означает, что условное распределение ξ_1 при условии, что $\xi_2 = x_2$ является гауссовским с параметрами a и σ^2. Отсюда следует

$$
\begin{aligned}
x_1 = g(x_2) = E(\xi_1|\xi_2 = x_2) = &\int_{-\infty}^{+\infty} x_1 f_{\xi_1|\xi_2}(x_1|\xi_2 = x_2)\,dx_1 = \\
&= a = a_1 + \rho\frac{\sigma_1}{\sigma_2}(x_2 - a_2)\,.
\end{aligned}
$$

Таким образом, кривая среднеквадратической регрессии ξ_1 на ξ_2 совпадает с прямой линейной среднеквадратической регрессии ξ_1 на ξ_2 (см. формулу (7.20)). Зависимость формы эллипса рассеивания от коэффициента корреляции ρ иллюстрируют рис. 3.1 и 3.2. Линия среднеквадратической регрессии ординаты ξ_2 на абсциссу ξ_1 проходит через верхнюю и нижнюю точки эллипса и делит пополам отрезок любой горизонтальной прямой, концы которого лежат на эллипсе.

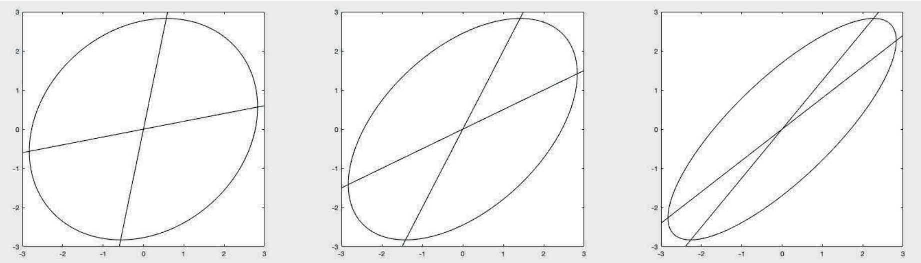


Рис. 3.1. Эллипсы рассеивания и линии среднеквадратической регрессии компонент двумерного гауссовского случайного вектора друг на друга для плотности распределения вероятностей (С.4), где $a_1 = a_2 = 0$, $\sigma_1 = \sigma_2 = 1$, а ρ принимает значения $0,2$; $0,5$; $0,8$ соответственно

Линия среднеквадратической регрессии абсциссы ξ_1 на ординату ξ_2 проходит через крайнюю правую и крайнюю левую точки эллипса и делит пополам отрезок любой вертикальной прямой, концы которого лежат на эллипсе. Видно, что при стремлении ρ к нулю эллипс приближается к кругу, а при стремлении $|\rho|$ к единице он вытягивается. В пределе при $|\rho| \to 1$ линии регрессии сливаются в прямую с угловым коэффициентом $\pm\dfrac{\sigma_2}{\sigma_1}$, а эллипс рассеивания сжимается в отрезок этой прямой. В результате ξ_2 становится линейной функцией от ξ_1.

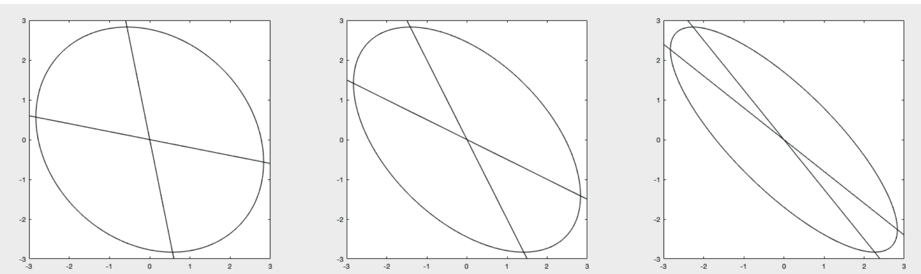


Рис. 3.2. Эллипсы рассеивания и линии среднеквадратической регрессии компонент двумерного гауссовского случайного вектора друг на друга для плотности распределения вероятностей (С.4), где $a_1 = a_2 = 0$, $\sigma_1 = \sigma_2 = 1$, а ρ принимает значения $-0,2$; $-0,5$; $-0,8$ соответственно

С.3. Характеристическая функция многомерного нормального распределения

Определение С.3.1. Характеристическая функция n-мерного случайного вектора $\vec{\xi}$ определяется равенством

$$\varphi_{\vec{\xi}}(\overline{t}) = Ee^{i(\vec{\xi},\overline{t})}, \quad \overline{t} \in \mathbb{R}^n .$$

Так же как и в одномерном случае, использование характеристических функций опирается на взаимную однозначность соответствия между характеристическими функциями и многомерными распределениями вероятностей. Аналогично характеристической функции одномерной случайной величины (см. свойство (φ-I) на стр. 143) $\varphi_{\vec{\xi}}(\overline{t})$ определена для любого $\overline{t} \in \mathbb{R}^n$ и удовлетворяет неравенству $|\varphi_{\vec{\xi}}(\overline{t})| \le 1$, при этом $\varphi_{\vec{\xi}}(\overline{0}) = 1$.

Следующая формула является многомерным аналогом свойства (φ-II) характеристических функций.

$$\boxed{\varphi_{B\vec{\xi}+\overline{a}}(\overline{t}) = e^{i(\overline{a},\overline{t})}\varphi_{\vec{\xi}}\left(B^T\overline{t}\right) .} \tag{C.5}$$

Доказательство. Действительно,

$$\varphi_{B\vec{\xi}+\overline{a}}(\overline{t}) = Ee^{i(B\vec{\xi}+\overline{a},\overline{t})} = e^{i(\overline{a},\overline{t})}Ee^{i(\vec{\xi},B^T\overline{t})} = e^{i(\overline{a},\overline{t})}\varphi_{\vec{\xi}}\left(B^T\overline{t}\right) . \qquad \blacksquare$$

Пусть $\vec{\eta} = (\eta_1, \eta_2, \ldots, \eta_n) \sim N(\overline{0}, E)$, т. е. $\vec{\eta}$ — n-мерный сферически-симметрический гауссовский случайный вектор. Тогда, пользуясь независимостью его компонент, получаем:

$$\begin{aligned}\varphi_{\vec{\eta}}(\overline{t}) = Ee^{i(\vec{\eta},\overline{t})} = Ee^{i\sum\limits_{j=1}^{n}\eta_j t_j} = E\prod_{j=1}^{n} e^{i\eta_j t_j} = \prod_{j=1}^{n} Ee^{i\eta_j t_j} = \\ = \prod_{j=1}^{n}\varphi_{\eta_j}(t_j) = \prod_{j=1}^{n} e^{-\frac{t_j^2}{2}} = e^{-\frac{1}{2}\sum\limits_{j=1}^{n}\frac{t_j^2}{2}} = e^{-\frac{1}{2}(\overline{t},\overline{t})} = e^{-\frac{1}{2}\|\overline{t}\|^2} .\end{aligned} \tag{C.6}$$

Пусть теперь $\vec{\xi} \sim N(\overline{a}, \Lambda)$ — невырожденный n-мерный гауссовский случайный вектор. Тогда, по следствию С.1.2, он представим в виде $\vec{\xi} = B\vec{\eta} + \overline{a}$, где B — невырожденная матрица $n \times n$, такая, что

$\Lambda = BB^T$, а $\vec{\eta}$ — сферически-симметрический гауссовский случайный вектор. По формуле (С.5), используя равенство (С.6), получаем

$$\varphi_{\vec{\xi}}(\overline{t}) = e^{i(\overline{a},\overline{t})}\varphi_{\vec{\eta}}\left(B^T\overline{t}\right) = e^{i(\overline{a},\overline{t})-\frac{1}{2}\left(B^T\overline{t},B^T\overline{t}\right)} = \\ = e^{i(\overline{a},\overline{t})-\frac{1}{2}\left(BB^T\overline{t},\overline{t}\right)} = e^{i(\overline{a},\overline{t})-\frac{1}{2}\left(\Lambda\overline{t},\overline{t}\right)}.$$

Таким образом, мы доказали следующее предложение:

Предложение С.3.1. Характеристическая функция невырожденного случайного вектора $\xi \sim N(\overline{a},\Lambda)$ имеет вид

$$\varphi_{\vec{\xi}}(\overline{t}) = e^{i(\overline{a},\overline{t})-\frac{1}{2}\left(\Lambda\overline{t},\overline{t}\right)}. \tag{С.7}$$

С.4. Вырожденное многомерное нормальное распределение

Оказывается, функция (С.7) является характеристической функцией некоторого распределения вероятностей и в том случае, когда матрица Λ является не положительно определенной, а лишь неотрицательно определенной, т. е. вырожденной. Соответствующее распределение вероятностей называют ***вырожденным многомерным гауссовским (нормальным)***.

Следующее предложение представляет собой аналог следствия С.1.2 для вырожденных гауссовских случайных векторов.

Предложение С.4.1. Пусть $\vec{\xi} \sim N(\overline{a},\Lambda)$ — n-мерный гауссовский случайный вектор, где Λ — неотрицательно-определенная вырожденная матрица. Тогда $\vec{\xi} = B\vec{\eta}+\overline{a}$, где B — матрица $n\times m$, $m<n$, а $\vec{\eta}$ — m-мерный сферически-симметрический гауссовский случайный вектор.

Доказательство. В условиях предложения С.4.1 существует невырожденное ортогональное преобразование T, приводящее матрицу Λ к диагональному виду,

$$\Theta = \operatorname{diag}(\theta_1,\theta_2,\ldots,\theta_n),$$

где $\theta_i, i = 1, \ldots, n$ — собственные значения Λ. Поскольку она вырождена, среди них есть нули. Пусть, без ограничения общности, $\theta_1 \neq 0, \ldots, \theta_m \neq 0$, а $\theta_{m+1} = \cdots = \theta_n = 0$. Запишем T и Θ в блочном виде:

$$T = (T_1 \; T_0)\ , \quad \Theta = \begin{pmatrix} \Theta_1 & 0 \\ 0 & 0 \end{pmatrix}\ ,$$

где T_1 — матрица $n \times m$, столбцы которой являются координатами собственных векторов матрицы Λ, отвечающих ненулевым собственным значениям $\theta_1, \ldots, \theta_m$; T_0 — матрица $n \times (n-m)$, столбцы которой являются координатами собственных векторов матрицы Λ, соответствующих нулевым собственным значениям; $\Theta_1 = \mathrm{diag}(\theta_1, \ldots, \theta_m)$. Имеем:

$$\Lambda = T\Theta T^T = (T_1 \; T_0) \begin{pmatrix} \Theta_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} T_1^T \\ T_0^T \end{pmatrix} =$$
$$= (T_1\Theta_1 \;\; 0) \begin{pmatrix} T_1^T \\ T_0^T \end{pmatrix} = T_1\Theta_1 T_1^T\ .$$

Положим $B = T_1\sqrt{\Theta_1}$. Это матрица $n \times m$ и $\Lambda = BB^T$. Пусть $\vec{\eta}$ — m-мерный сферически-симметрический гауссовский случайный вектор. Тогда компоненты случайного вектора $\vec{\xi} = B\vec{\eta} + \overline{a}$, представляющие собой линейные функции от гауссовских случайных величин, являются гауссовскими случайными величинами, при этом $E\vec{\xi} = \overline{a}$, а для ковариационных моментов имеем равенство

$$\mathrm{Cov}\,(\xi_i, \xi_j) = \mathrm{Cov}\left(\sum_{k=1}^{m} b_{ik}\eta_k + a_i, \sum_{l=1}^{m} b_{jl}\eta_l + a_j\right) =$$
$$= \sum_{k=1}^{m}\sum_{l=1}^{m} b_{ik}b_{jl}\mathrm{Cov}\ (\eta_k, \eta_l) = \sum_{k=1}^{m} b_{ik}b_{jk} = \lambda_{ij}\ .$$

Таким образом, $\vec{\xi}$ распределен по закону $N(\overline{a}, \Lambda)$. ■

R Из доказательства предложения C.4.1 видно, что носителем вырожденного n-мерного нормального распределения вероятностей является m-мерное гиперпространство $L \subset \mathbb{R}^n$, вида

$$L = \left\{\overline{x} = \overline{a} + \alpha_1\overline{t}_1 + \ldots \alpha_m\overline{t}_m \in \mathbb{R}^n \mid \alpha_i \in \mathbb{R}\, , i = 1, \ldots, m\right\}\ ,$$

где $\overline{t}_i$, $i = 1, \ldots, m$ — собственные векторы ковариационной матрицы Λ этого распределения вероятностей, отвечающие ее ненулевым собственным значениям.

■ Пример С.4.1. Пусть

$$\vec{\xi} = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \sim N(\overline{a}, \Lambda), \text{ где } \overline{a} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \ \Lambda = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Матрица Λ имеет собственные значения $\lambda_1 = 2$ и $\lambda_2 = 0$. Векторы $\overline{t}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$ и $\overline{t}_2 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$ — единичные собственные векторы, соответствующие этим собственным значениям. Таким образом,

$$B = T_1\sqrt{\Theta_1} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}\sqrt{2} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \vec{\xi} = B\eta + \overline{a} = \begin{pmatrix} \eta + 2 \\ \eta + 3 \end{pmatrix},$$

а носителем данного распределения вероятностей является прямая, проходящая через точку $\overline{a}$ с направляющим вектором $\overline{t}_1$. ■

С.5. Задачи для самостоятельного решения

1. Случайный вектор $\vec{\xi} = (\xi_1, \xi_2)$ имеет нормальное распределение с плотностью

$$f_{\vec{\xi}}(x, y) = \frac{1}{6\pi} \exp\left[-\frac{1}{18}\Big(5(x-1)^2 - 2(x-1)y + 2y^2\Big)\right].$$

Найти:

a) математическое ожидание и ковариационную матрицу $\vec{\xi}$;

b) плотности распределения $f_{\xi_1}(x)$ и $f_{\xi_2}(x)$;

c) $\mathrm{P}\,(2\xi_1 - 3\xi_2 < -4)$;

d) $\mathrm{P}\,(\xi_1 - 1 < \xi_2 < -3\xi_1 + 3)$;

e) Построить эллипс рассеивания и кривые среднеквадратической регрессии компонент случайного вектора друг на друга.

2. Случайный вектор $\vec{\xi} = (\xi_1, \xi_2)$ имеет нормальное распределение с плотностью

$$f_{\vec{\xi}}(x, y) = \frac{1}{2\pi\sqrt{3}} \exp\left[-\frac{1}{6}\Big(4(x-1)^2 + 2(x-1)(y+1) + (y+1)^2\Big)\right].$$

Найти:

a) математическое ожидание и ковариационную матрицу $\bar{\xi}$;

b) плотности распределения $f_{\xi_1}(x)$ и $f_{\xi_2}(x)$;

c) $\mathrm{P}\left(-\xi_1 + 3\xi_2 < 2\right)$;

d) $\mathrm{P}\left(2\xi_1 - 3 < \xi_2 < 3\xi_1 - 4\right)$;

e) Построить эллипс рассеивания и кривые среднеквадратической регрессии компонент случайного вектора друг на друга.

Приложение D. Распределение χ^2 и теорема Пирсона

D.1. Распределение χ^2

Определение D.1.1. Пусть $\vec{\eta} \sim N(\overline{0}, E)$ — n-мерный случайный вектор, $\xi = \sum\limits_{k=1}^{n} \eta_k^2$. Распределение вероятностей ξ называется распределением ***хи-квадрат с n степенями свободы***. При этом пишут $\xi \sim \chi_n^2$.

Найдем плотность распределения вероятностей хи-квадрат.

Предложение D.1.1. Пусть $\eta \sim N(0,1)$, $\xi = \eta^2$. Тогда $\xi \sim \Gamma\left(\frac{1}{2}, 2\right)$.

Доказательство. Носителем распределения вероятностей случайной величины ξ является промежуток $[0;+\infty)$, так как $\xi \geq 0$. Пусть $x \geq 0$, тогда

$$F_\xi(x) = P\left(\xi < x\right) = P\left(\eta^2 < x\right) = P\left(-\sqrt{x} < \eta < \sqrt{x}\right) =$$
$$= \frac{1}{\sqrt{2\pi}} \int\limits_{-\sqrt{x}}^{\sqrt{x}} e^{-\frac{t^2}{2}} dt = \sqrt{\frac{2}{\pi}} \int\limits_{0}^{\sqrt{x}} e^{-\frac{t^2}{2}} dt\,.$$

Дифференцируя, получаем:

$$f_\xi(x) = \frac{d}{dx} F_\xi(x) = \frac{1}{\sqrt{2\pi}} \frac{e^{-\frac{x}{2}}}{\sqrt{x}} = \frac{x^{k-1}}{\theta^k \Gamma(k)} e^{-\frac{x}{\theta}}\,,\ \text{где } k = \frac{1}{2}\,,\ \theta = 2\,. \quad \blacksquare$$

Предложение D.1.2. Распределение вероятностей χ_n^2 совпадает с распределением $\Gamma\left(\frac{n}{2}, 2\right)$, т. е. плотность распределения χ_n^2 имеет вид:

$$f_{\chi_n^2}(x) = \frac{x^{\frac{n}{2}-1}}{2^n\Gamma\left(\frac{n}{2}\right)} e^{-\frac{x}{2}} \cdot \mathbf{1}_{[0;+\infty)}(x)\,.$$

Доказательство. При $n = 1$ утверждение уже доказано. Пусть $n > 1$, $\xi \sim \chi_n^2$, т. е. $\xi = \sum\limits_{k=1}^{n} \eta_k^2$, где случайные величины $\eta_k \sim N(0,1)$ независимы. Поскольку $\eta_k^2 \sim \Gamma\left(\frac{1}{2}, 2\right)$, характеристическая функция этой случайной величины имеет вид

$$\varphi_{\eta_k^2}(t) = \frac{1}{\sqrt{1-2it}}$$

(см. пример 9.2.4, формула (9.3)). Отсюда, по свойству (φ-V) характеристических функций, в силу независимости случайных величин $\eta_1^2, \ldots, \eta_n^2$, получаем:

$$\varphi_\xi(t) = \varphi_{\eta_1^2+\cdots+\eta_n^2}(t) = \prod_{k=1}^{n} \varphi_{\eta_k^2}(t) = \frac{1}{\sqrt{(1-2it)^n}}\,.$$

Это характеристическая функция распределения $\Gamma\left(\frac{n}{2}, 2\right)$. ∎

D.2. Теорема Пирсона

Пусть выборка $\vec{\xi}_n$ соответсвует случайной величине ξ, $\{B_k\}_{k=1}^r$ — набор непересекающихся промежутков в $\mathbb{R}$ таких, что для любого k $p_k := P\big(\xi \in B_k\big) \neq 0$ и $\left\{\xi \in \bigsqcup\limits_{k=1}^{r} B_k\right\} = \Omega$, так что в результате $\sum\limits_{k=1}^{r} p_k = 1$. Пусть $\nu_k = \sum\limits_{i=1}^{n} \mathbf{1}_{B_k}(\xi_i)$.

Теорема D.2.1. — Пирсон. Распределение вероятностей суммы $\sum\limits_{k=1}^{r} \frac{(\nu_k - np_k)^2}{np_k}$ при $n \to \infty$ сходится к распределению χ^2_{r-1}.

Доказательство. Рассмотрим случайный вектор

$$\vec{\theta} = (\theta_1, \ldots, \theta_r)^T = \left(\frac{\nu_1 - np_1}{\sqrt{np_1}}, \ldots, \frac{\nu_r - np_r}{\sqrt{np_r}} \right)^T .$$

Построим его характеристическую функцию:

$$\varphi_{\vec{\theta}}(\bar{t}) = Ee^{i(\vec{\theta}, \bar{t})} = Ee^{i \sum\limits_{k=1}^{r} \frac{\nu_k - np_k}{\sqrt{np_k}} t_k} = Ee^{i \sum\limits_{k=1}^{r} \nu_k \frac{t_k}{\sqrt{np_k}} - i \sum\limits_{k=1}^{r} t_k \sqrt{np_k}} =$$
$$= e^{-i \sum\limits_{k=1}^{r} t_k \sqrt{np_k}} Ee^{i \sum\limits_{k=1}^{r} \sum\limits_{j=1}^{n} \mathbf{1}_{B_k}(\xi_j) \frac{t_k}{\sqrt{np_k}}} .$$

Для вычисления математического ожидания поменяем порядок суммирования в аргументе экспоненты и преобразуем ее в произведение:

$$Ee^{i \sum\limits_{k=1}^{r} \sum\limits_{j=1}^{n} \mathbf{1}_{B_k}(\xi_j) \frac{t_k}{\sqrt{np_k}}} = Ee^{i \sum\limits_{j=1}^{n} \sum\limits_{k=1}^{r} \mathbf{1}_{B_k}(\xi_j) \frac{t_k}{\sqrt{np_k}}} =$$
$$E \prod_{j=1}^{n} e^{i \sum\limits_{k=1}^{r} \mathbf{1}_{B_k}(\xi_j) \frac{t_k}{\sqrt{np_k}}} = \left(Ee^{i \sum\limits_{k=1}^{r} \mathbf{1}_{B_k}(\xi_j) \frac{t_k}{\sqrt{np_k}}} \right)^n .$$

Здесь мы воспользовались тем, что множители $e^{i \sum\limits_{k=1}^{r} \mathbf{1}_{B_k}(\xi_j) \frac{t_k}{\sqrt{np_k}}}$ независимы, так как представляют собой функции от независимых случайных величин ξ_j, и одинаково распределены. Поскольку ξ_j попадает только в один из промежутков B_k и $P\big(\xi_j \in B_k\big) = p_k$, случайная величина $e^{i \sum\limits_{k=1}^{r} \mathbf{1}_{B_k}(\xi_j) \frac{t_k}{\sqrt{np_k}}}$ принимает значения $e^{i \frac{t_k}{\sqrt{np_k}}}$ с вероятностями p_k. Следовательно, получаем

$$Ee^{i \sum\limits_{k=1}^{r} \sum\limits_{j=1}^{n} \mathbf{1}_{B_k}(\xi_j) \frac{t_k}{\sqrt{np_k}}} = \sum_{k=1}^{r} e^{i \frac{t_k}{\sqrt{np_k}}} p_k \, .$$

Таким образом,

$$\varphi_{\vec{\theta}}(\bar{t}) = e^{-i \sum\limits_{k=1}^{r} t_k \sqrt{np_k}} \left(\sum_{k=1}^{r} e^{i \frac{t_k}{\sqrt{np_k}}} p_k \right)^n .$$

Прологарифмируем $\varphi_{\vec{\theta}}(\overline{t})$. Используя формулу Тейлора–Маклорена сначала для экспоненты, а затем для логарифма, получим:

$$\ln \varphi_{\vec{\theta}}(\overline{t}) = -i \sum_{k=1}^{r} t_k \sqrt{np_k} + n \ln \left(\sum_{k=1}^{r} e^{i \frac{t_k}{\sqrt{np_k}}} p_k \right) =$$

$$= -i \sum_{k=1}^{r} t_k \sqrt{np_k} + n \ln \left(\sum_{k=1}^{r} p_k \left(i \frac{t_k}{\sqrt{np_k}} - \frac{t_k^2}{2np_k} + o\left(\frac{1}{n}\right) \right) \right) =$$

$$= -i \sum_{k=1}^{r} t_k \sqrt{np_k} + n \ln \left(i \sum_{k=1}^{r} t_k \sqrt{\frac{p_k}{n}} - \frac{1}{2n} \sum_{k=1}^{r} t_k^2 + o\left(\frac{1}{n}\right) \right) =$$

$$= -i \sum_{k=1}^{r} t_k \sqrt{np_k} + n \left(i \sum_{k=1}^{r} t_k \sqrt{\frac{p_k}{n}} - \frac{1}{2n} \sum_{k=1}^{r} t_k^2 + \frac{1}{2n} \left(\sum_{k=1}^{r} t_k \sqrt{p_k} \right)^2 + \right.$$

$$\left. + o\left(\frac{1}{n}\right) \right) = -\frac{1}{2} \sum_{k=1}^{r} t_k^2 + \frac{1}{2} \left(\sum_{k=1}^{r} t_k \sqrt{p_k} \right)^2 + o(1)\,.$$

Отсюда следует сходимость

$$\varphi_{\vec{\theta}}(\overline{t}) \xrightarrow[n\to\infty]{} \exp \left[-\frac{1}{2} \left(\sum_{k=1}^{r} t_k^2 - \left(\sum_{k=1}^{r} t_k \sqrt{p_k} \right)^2 \right) \right] =$$

$$= \exp \left[-\frac{1}{2} \left(\Lambda \overline{t}, \overline{t} \right) \right] =: \varphi_{\vec{\zeta}}(\overline{t})\,,$$

где

$$\Lambda = \begin{pmatrix} 1-p_1 & -\sqrt{p_1p_2} & -\sqrt{p_1p_2} & \ldots & -\sqrt{p_1p_r} \\ -\sqrt{p_2p_1} & 1-p_2 & -\sqrt{p_2p_3} & \ldots & -\sqrt{p_2p_r} \\ -\sqrt{p_3p_1} & -\sqrt{p_3p_2} & 1-p_3 & \ldots & -\sqrt{p_3p_r} \\ & \vdots & & \ldots & \vdots \\ -\sqrt{p_rp_1} & -\sqrt{p_rp_2} & -\sqrt{p_rp_3} & \ldots & 1-p_r \end{pmatrix}.$$

Функция $\varphi_{\vec{\zeta}}(\overline{t})$ является характеристической функцией многомерного нормального распределения с математическим ожиданием $\overline{0}$ и ковариационной матрицей Λ (см. предложение C.3.1). По теореме непрерывности (теорема 10.4.1) из этого следует, что при $n \to \infty$ случайный вектор $\vec{\theta}$ сходится по распределению к $\vec{\zeta} \sim N(\overline{0}, \Lambda)$.

Сделаем в функции $\varphi_{\vec{\zeta}}(\overline{t})$ замену переменных $\overline{u} = B\overline{t}$, где B — некоторая ортонормированная матрица, т. е. матрица, строки которой образуют ортонормированный базис пространства $\mathbb{R}^n$. При этом потребуем, чтобы выполнялось равенство

$$u_r = \sum_{k=1}^{r} t_k \sqrt{p_k}\,, \tag{D.1}$$

т. е. чтобы последняя строка матрицы B имела вид $(\sqrt{p_1}, \ldots, \sqrt{p_r})$. Ортонормированность матрицы приводит к тому, что $BB^T = E$, а значит $B^T = B^{-1}$ и $\overline{t} = B^T\overline{u}$. При этом

$$\sum_{k=1}^{r} t_k^2 = (\overline{t}, \overline{t}) = (B^T\overline{u}, B^T\overline{u}) = (\overline{u}, BB^T\overline{u}) = (\overline{u}, \overline{u}) = \sum_{k=1}^{r} u_k^2\,. \tag{D.2}$$

В результате замены, в силу равенств (D.1) и (D.2), получаем

$$\varphi_{\vec{\zeta}}(\overline{t}) = \exp\left[-\frac{1}{2}\left(\sum_{k=1}^{r} t_k^2 - \left(\sum_{k=1}^{r} t_k\sqrt{p_k}\right)^2\right)\right] =$$
$$= \exp\left[-\frac{1}{2}\left(\sum_{k=1}^{r} u_k^2 - u_r^2\right)\right] = \exp\left[-\frac{1}{2}\sum_{k=1}^{r-1} u_k^2\right]\,.$$

Но по формуле (C.5) имеем:

$$\varphi_{\vec{\zeta}}(\overline{t}) = \varphi_{\vec{\zeta}}(B^T\overline{u}) = \varphi_{B\vec{\zeta}}(\overline{u})\,.$$

Это означает, что случайный вектор $B\vec{\zeta}$, получающийся из $\vec{\zeta}$ ортонормированным преобразованием (поворотом вокруг $\overline{0}$), имеет $r-1$-мерное сферически-симметрическое гауссовское распределение вероятностей. Следовательно, распределение вероятностей самого вектора $\vec{\zeta}$ также является вырожденным сферически-симметрическим гауссовским, сосредоточенным на некотором $r-1$-мерном подпространстве пространства $\mathbb{R}^r$.

Распределение вероятностей суммы $\displaystyle\sum_{k=1}^{r}\frac{(\nu_k - np_k)^2}{np_k} = \left|\vec{\theta}\right|^2$ сходится к распределению вероятностей случайной величины $\left|\vec{\zeta}\right|^2$. Из доказанного выше следут, что оно является распределением хи-квадрат с $r-1$ степенями свободы. ■

Приложение Е. Задания лабораторных работ по математической статистике

Е.1. Лабораторная работа №1 «Оценки параметров распределения вероятностей по выборке»

Эмпирическая функция распределения и эмпирическая плотность

1. Используя генераторы псевдослучайных чисел *MatLab* (функция `random`), сгенерировать реализацию выборки, соответствующей некоторому распределению вероятностей, задав тип распределения вероятностей (гамма-распределение, нормальное, равномерное и т. п.), его параметры и n – объем выборки.
2. Для сгенерированной реализации выборки построить эмпирическую функцию распределения $F_n(x)$. Для этого упорядочить сгенерированную реализацию с помощью функции `sort`. Далее для построения эмпирической функции распределения использовать функцию `stairs`. Сравнить график построенной эмпирической функции распределения с графиком теоретической функции распределения, построив эти графики в одном окне. Для вычисления значений теоретической функции распределения использовать функцию *MatLab* вида `xxxxcdf`, где префикс `xxxx` задает тип распределения вероятностей. Например, функция `normcdf(x,a,sigma)` вычисляет значения в точках, содержащихся в массиве `x`, функции распределения, соответствующей

нормальному распределению с параметрами `a` и `sigma`. Продемонстрировать, что $F_n(x)$ — случайная величина, перезапуская скрипт, тем самым генерируя новые реализации выборки. Придавая n значения $10, 20, 40, 100, 1000$, проиллюстрировать теорему Гливенко о равномерной сходимости эмпирических функций распределения при стремлении к бесконечности объема выборки.

3. Построить гистограмму сгенерированной реализации выборки, используя функцию Matlab `histogram`. Построенная гистограмма представляет собой график эмпирической плотности распределения, которому соответствует выборка, масштабированный по оси ординат с коэффициентом nh, где h — шаг разбиения, по которому строится гистограмма. Сравнить график эмпирической плотности распределения с графиком теоретической плотности, построив масштабированный по оси ординат с коэффициентом nh график теоретической плотности распределения в одном окне с гистограммой. Для вычисления значений теоретической плотности распределения использовать функцию *MatLab* вида `xxxxpdf`, где префикс `xxxx` задает тип распределения вероятностей. Например, функция `normpdf(x,a,sigma)` вычисляет значения в точках, содержащихся в массиве `x`, плотности нормального распределения вероятностей с параметрами `a` и `sigma`. Проиллюстрировать теорему о сходимости эмпирической плотности распределения, меняя n так же, как в п. 2. Показать, что сходимость имеет место только при согласовании с n количества промежутков, по которым строится гистограмма (количество промежутков определяется параметром h в формулировке теоремы).
4. С помощью построенного скрипта провести численные эксперименты, меняя тип распределения вероятностей (всего рассмотреть три разных типа).

Точечные и интервальные оценки моментов случайной величины по выборке

1. Используя генераторы псевдослучайных чисел *MatLab*, сгенерировать m реализаций выборки, соответствующей некоторому распределению вероятностей, задав тип распределения вероятностей (гамма-распределение, нормальное, равномерное и т. п.), его параметры и n – объем выборки.
2. Для каждой реализации выборки вычислить выборочное среднее, выборочный момент второго порядка, выборочный центральный момент второго порядка. Вывести графики изменения этих оценок в три окна. В те же окна вывести точные значения математического ожидания, второго момента и дисперсии, показать, как при повторении статистического эксперимента точечные оценки колеблются около точных значений оцениваемых параметров.
3. Для каждой реализации выборки вычислить границы доверительного интервала для математического ожидания, соответствующего доверительной вероятности $\alpha = 0,9$, используя известную дисперсию. Построить в отдельном окне графики левого и правого концов доверительного интервала и точного значения математического ожидания. Провести численные эксперименты, иллюстрирующие изменение размеров доверительного интервала и надежности покрытия доверительным интервалом оцениваемого математического ожидания при изменении объема выборки ($n = 50, 100, 1000, 10000$) и доверительной вероятности ($\alpha = 0,8,\ 0,9,\ 0,95,\ 0,99$). Как меняется поведение доверительного интервала, если известную дисперсию заменить на выборочную?

Е.2. Лабораторная работа №2 «Исследование распределения вероятностей по реализации выборки. Критерии согласия»

Исследовать две случайные величины по их выборкам. Реализации выборок даны в файлах `v_xx_y.mat`, где `xx` — номер варианта, `y` — номер задания. Для каждой реализации выборки в вычислительной среде *MatLab*, используя приемы, изученные в ходе выполнения лабораторной работы №1, выполнить следующее:

1. Исследовать реализацию выборки, определив ее объем, построив по ней эмпирическую функцию распределения, эмпирическую плотность распределения, вычислив выборочное среднее и выборочную дисперсию. Сформулировать нуль-гипотезу о характере распределения вероятностей, которому соответствует данная реализация выборки.
2. Сделать субъективную визуальную оценку нуль-гипотезы, построив в одном окне графики эмпирической функции распределения и гипотетической функции распределения, соответствующей нуль-гипотезе, а в другом окне — графики эмпирической и гипотетической плотностей распределения вероятностей.
3. Для объективной оценки нуль-гипотезы вычислить статистику Колмогорова, применить критерий Колмогорова с уровнем значимости 0,05. По значению статистики Колмогорова вычислить критический уровень значимости для данной реализации выборки, дать его вероятностную интерпретацию.
4. Варьируя параметры распределения вероятностей, соответствующего нуль-гипотезе, минимизировать, если это возможно, значение статистики Колмогорова. Чему равен критический уровень значимости для этого значения статистики? Сформулировать вероятностный смысл полученного значения критического уровня значимости.
5. Вычислить статистику Пирсона, используя для ее построения интервалы, по которым построена гистограмма реализации вы-

борки. Применить критерий Пирсона с уровнем значимости 0,05, по значению статистики Пирсона вычислить критический уровень значимости.

6. Варьируя параметры распределения вероятностей, соответствующего нуль-гипотезе, минимизировать, если это возможно, значение статистики Пирсона. Чему равен критический уровень значимости для этого значения статистики? Сформулировать вероятностный смысл полученного значения критического уровня значимости.

Приложение F. Ответы к задачам

1.6.

1. $\dfrac{24}{10!} = \dfrac{1}{151200}$.

2. $\dfrac{M}{N}$.

3. $\dfrac{2(n!)^2}{(2n)!}$.

4. $\dfrac{28 \cdot 21 + 20 \cdot 23 + 12 \cdot 25 + 4 \cdot 27}{64 \cdot 63} = \dfrac{13}{36}$.

5. $\dfrac{21 \cdot 12 + 7 \cdot 6}{28 \cdot 27} = \dfrac{7}{18}$.

6. $\dfrac{1}{2} + \dfrac{b}{24}$ при $0 < b \leq 4$, $1 - \dfrac{2}{3\sqrt{b}}$ при $b > 4$.

7. $\dfrac{1}{4}$.

8. a) $\left(1 - \dfrac{r}{R}\right)^2$; b) $1 - \dfrac{4}{\pi}\arcsin\dfrac{r}{2R} - \dfrac{r}{\pi R^2}\sqrt{4R^2 - r^2} + \dfrac{r^2}{\pi R^2}$.

9. $\dfrac{1}{\pi}\operatorname{arctg}\dfrac{h}{l}$.

2.4.

1. a) $0,622$; b) $0,166$.

2. $p_1 = 0,4$, $p_2 = 0,3$, $p_3 = 0,2$, $p_4 = 0,1$.

3. a) $\dfrac{k-1}{20}$; b) $\dfrac{20!}{20^k(20-k)!}$.

4. с сильным игроком.

5. $\dfrac{1}{3}$.

6. $p_0 = \dfrac{1}{19}$, $p_1 = p_2 = \dots = p_9 = \dfrac{2}{19}$, $p_{10} = p_{11} = \dots = p_{18} = 0$.

8. a) да; b) нет.

9. a) Вероятность выигрыша первого — $\dfrac{2}{3}$, второго — $\dfrac{1}{3}$; b) Вероятность выигрыша первого — $\dfrac{4}{7}$, второго — $\dfrac{2}{7}$, третьего — $\dfrac{1}{7}$.

3.3.

1. $\dfrac{38}{105}$.

2. $\dfrac{75}{392}$.

3. $\dfrac{19}{42}$.

4. Вероятности того, что не попал 1-й, 2-й и 3-й стрелки равны $\dfrac{3}{47}$, $\dfrac{8}{47}$ и $\dfrac{36}{47}$ соответственно.

5. $\dfrac{8}{9}$.

6. $\dfrac{1}{12}$.

4.3.

1. Если $np - q \in \mathbb{Z}$ (что эквивалентно условию $(n+1)p \in \mathbb{Z}$), то $P_n(np-q) = P_n((n+1)p)$ — максимальное значение вероятности $P_n(m)$ (существует два наиболее вероятных значения числа успехов в серии из n испытаний Бернулли с вероятностью успеха p); если

$np - q \notin \mathbb{Z}$, то наиболее вероятное значение числа успехов единственно и равно $[(n+1)p]$; $m_0 = 2$, $P_6(1,3) = 0,774144$.

2. $P_{800}(2, 800) \approx 0,475$.

3. от 226 до 274.

4. 77.

5. 649740.

6. $0,268$.

7. 9604.

5.6.

1. $$\sigma(\mathcal{K}) = \Big\{(-\infty;0), (-\infty;1), [0;+\infty), [1;+\infty), [0;1), (-\infty;0)\bigcup[1;+\infty), \varnothing, \mathbb{R}\Big\}.$$

2.

x_i	1	2	3	4	5
p_i	1/25	3/25	5/25	7/25	9/25

3.

x_i	0	1	2	3
p_i	0,192	0,464	0,296	0,048

, $P\big(\xi \geq 2\big) = 0,344$.

4. $P(\xi = m + k) = C_{m+k-1}^{m-1} p^m q^k$, $k = 0, 1, 2, \ldots$.

5. $F_\xi(x) = \dfrac{x^2}{2} \cdot \mathbf{1}_{[0;1]}(x) + \left(1 - \dfrac{(2-x)^2}{2}\right) \cdot \mathbf{1}_{(1;2]}(x) + \mathbf{1}_{(2;+\infty)}(x)$;

$f_\xi(x) = x \cdot \mathbf{1}_{[0;1]}(x) + (2 - x) \cdot \mathbf{1}_{(1;2]}(x)$;

$F_\eta(x) = \dfrac{(x+1)^2}{2} \cdot \mathbf{1}_{[-1;0]}(x) + \left(1 - \dfrac{(1-x)^2}{2}\right) \cdot \mathbf{1}_{(0;1]}(x) + \mathbf{1}_{(1;+\infty)}(x)$;

$f_\eta(x) = (x + 1) \cdot \mathbf{1}_{[-1;0]}(x) + (1 - x) \cdot \mathbf{1}_{(0;1]}(x)$.

6. $F_\varphi(x) = \dfrac{1}{2}\operatorname{tg} x \cdot \mathbf{1}_{[0;\frac{\pi}{4}]}(x) + \left(1 - \dfrac{1}{2}\operatorname{ctg} x\right) \cdot \mathbf{1}_{\left(\frac{\pi}{4};\frac{\pi}{2}\right]}(x) + \mathbf{1}_{\left(\frac{\pi}{2};+\infty\right)}(x);$

$f_\varphi(x) = \dfrac{1}{2\cos^2 x} \cdot \mathbf{1}_{[0;\frac{\pi}{4}]}(x) + \dfrac{1}{2\sin^2 x} \cdot \mathbf{1}_{\left(\frac{\pi}{4};\frac{\pi}{2}\right]}(x);$

$F_\rho(x) = \dfrac{\pi x^2}{4} \cdot \mathbf{1}_{[0;1]}(x) + \left(\sqrt{x^2-1} + x^2\left(\dfrac{\pi}{4} - \arccos\dfrac{1}{x}\right)\right) \cdot \mathbf{1}_{(1;\sqrt{2}]}(x) +$

$+\mathbf{1}_{(\sqrt{2};+\infty)}(x);$

$f_\rho(x) = \dfrac{\pi x}{2} \cdot \mathbf{1}_{[0;1]}(x) + \left(\dfrac{\pi x}{2} - 2x\arccos\dfrac{1}{x}\right) \cdot \mathbf{1}_{(1;\sqrt{2}]}(x);$

$P\left(\dfrac{\pi}{6} < \varphi < \dfrac{\pi}{3}\right) = 1 - \dfrac{1}{\sqrt{3}} \approx 0,423;$

$P\left(\dfrac{1}{2} < \rho < \dfrac{2}{\sqrt{3}}\right) = \dfrac{1}{\sqrt{3}} + \dfrac{7\pi}{144} \approx 0,73.$

7. $F_\theta(x) = (x - x\ln x) \cdot \mathbf{1}_{[0;1]}(x) + \mathbf{1}_{(0;+\infty)}(x);\ \ f_\theta(x) = -\ln x;$

$P\left(\dfrac{1}{2} < \theta < \dfrac{3}{4}\right) = \dfrac{1}{4} + \ln\dfrac{2\sqrt[3]{3}}{3} \approx 0,211.$

8. a) $F_\eta(x) = (1 - e^{-\sqrt{x}}) \cdot \mathbf{1}_{[0;+\infty)}(x)$, $f_\eta(x) = \dfrac{1}{2\sqrt{x}} e^{-\sqrt{x}} \cdot \mathbf{1}_{[0;+\infty)}(x);$

b) $F_\eta(x) = \sqrt{x} \cdot \mathbf{1}_{[0;1]}(x) + \mathbf{1}_{(1;+\infty)}(x)$, $f_\eta(x) = \dfrac{1}{2\sqrt{x}} \cdot \mathbf{1}_{(0;1]}(x).$

9. $F_\eta(x) = \dfrac{2}{\pi}\arcsin x \cdot \mathbf{1}_{[0;1]}(x) + \mathbf{1}_{(1;+\infty)}(x);\ f_\eta(x) = \dfrac{2}{\pi\sqrt{1-x^2}} \cdot \mathbf{1}_{[0;1]}(x).$

6.6.

1.

$\eta \backslash \xi$	1	2	3	$\cdots$	n	$\cdots$
0	1/3	1/9	2/27	$\cdots$	$2^{n-2}/3^n$	$\cdots$
1	0	1/9	2/27	$\cdots$	$2^{n-2}/3^n$	$\cdots$

,

$\xi \sim \mathrm{Geom}\left(\dfrac{1}{3}\right)$,

η	0	1
p_i	2/3	1/3

, ξ и η зависимы т. к., например,

$P\big(\xi = 1, \eta = 1\big) = 0 \neq P(\xi = 1) \cdot P(\eta = 1) = \dfrac{1}{3} \cdot \dfrac{2}{3}.$

2. $p_{ijk} = P\big(\theta_1 = i, \theta_2 = j, \theta_3 = k\big) =$

$$= \begin{cases} 0\,, & i = j \vee j = k \vee k = i\,, \\ \dfrac{1}{N(N-1)(N-2)}\,, & i \neq j \wedge j \neq k \wedge k \neq i\,, \end{cases}$$

$p_i^1 = p_j^2 = p_k^3 = \dfrac{1}{N}$, случайные величины зависимы т. к., например, $p_{iii} = 0 \neq p_i^1 \cdot p_i^2 \cdot p_i^3 = \dfrac{1}{N^3}$.

3. $p_{ij} = P\big(\tau_2 = i, \tau_2 = j\big) = \dfrac{(N-1)(N-2)2^{j-1}}{N^{i+j}}\,, \tau_2 \sim \text{Geom}\left(\dfrac{N-1}{N}\right)\,,$

$\tau_3 \sim \text{Geom}\left(\dfrac{N-2}{N}\right)$, случайные величины независимы.

4. $f_{\varphi\rho}(x,y) = \dfrac{y}{\pi} \cdot \mathbf{1}_{[0;2\pi]\times[0;1]}(x,y)\,,\ f_{\varphi}(x) = \dfrac{1}{2\pi} \cdot \mathbf{1}_{[0;2\pi]}(x)\,,$

$f_{\rho}(x) = 2x \cdot \mathbf{1}_{[0;1]}(x)$, случайные величины независимы.

5. $f_{\varphi,\rho}(x,y) = \dfrac{1}{4\pi} \cdot e^{-\frac{y}{2}} \cdot \mathbf{1}_{[0;2\pi)\times[0;+\infty)}(x,y)\,,\ \varphi \sim U[0;2\pi)\,,$

$\theta \sim \text{Exp}\left(\dfrac{1}{2}\right)$.

6. $f_{\zeta,\nu}(x,y) = \dfrac{1}{2}e^{-x} \cdot \mathbf{1}_S(x,y)$, где $S = \{(x,y)\,|\,x \geq 0\,,\ -x \leq y \leq x\}\,,$

$f_{\zeta}(x) = xe^{-x} \cdot \mathbf{1}_{[0;+\infty)}(x)\,,\ (\zeta \sim \Gamma\,(2,1))\,,\ f_{\nu}(x) = \dfrac{1}{2}e^{-|x|}$, случайные величины зависимы.

7.5.

1. $E\xi = 3,8\,,\ D\xi = 1,36\,, \sigma_\xi \approx 1,166\,,$

$P\,(|\xi - E\xi| < \sigma_\xi) = P\big(\{\xi = 3\} + \{\xi = 4\}\big) = 0,48\,.$

2. $E\xi = 1,2\,,\ D\xi = 0,64\,,\ \sigma_\xi = 0,8\,,$ выполняется т. к.

$P\,(|\xi - E\xi| < 2\sigma_\xi) = 0,952\,.$

3. $E\xi = \dfrac{m}{p}\,,\ D\xi = \dfrac{m(1-p)}{p^2}\,.$

4. $E\xi = 1\,,\ E\eta = 0\,,\ D\xi = D\eta = \dfrac{1}{6}\,,$

$P\left(|\xi - E\xi| < \sigma_\xi\right) = P\left(|\eta - E\eta| < \sigma_\eta\right) = \dfrac{2\sqrt{6}-1}{6} \approx 0,65\,.$

5. $a = \dfrac{1}{8}\,,\ E\xi = 2,5\,,\ D\xi = 0,15\,,\ P\left(|\xi - E\xi| < 2\sigma_\xi\right) \approx 0,952\,.$

6. $a = \dfrac{3}{4}\,, E\xi = 0,8\,,\ D\xi = 0,16\,,\ P\left(|\xi - E\xi| < 2\sigma_\xi\right) = 0,9728\,.$

7. $\mathrm{Cov}\,(\zeta, \nu) = \rho_{\zeta\nu} = 0$, прямые линейной среднеквадратической регрессии: ν на ζ: $y = 0$, ζ на ν: $x = 0$.

8. $\mathrm{Cov}\,(\xi, \eta) = \dfrac{1}{40}\,,\quad \rho_{\xi\eta} = \dfrac{1}{2}\sqrt{\dfrac{35}{57}}\,,$ прямые линейной среднеквадратической регрессии: η на ξ: $y = \dfrac{2}{5} + \dfrac{8}{19}\left(x - \dfrac{5}{8}\right)$, ξ на η: $x = \dfrac{5}{8} + \dfrac{35}{96}\left(y - \dfrac{2}{5}\right)$.

9. $a = 3\,,\quad \mathrm{Cov}\,(\xi, \eta) = -\dfrac{3}{160}\,,\quad \rho_{\xi\eta} = -\sqrt{\dfrac{3}{19}}\,,$ прямые линейной среднеквадратической регрессии: η на ξ: $y = \dfrac{3}{8} - \dfrac{1}{2}\left(x - \dfrac{1}{4}\right)$, ξ на η: $x = \dfrac{1}{4} - \dfrac{6}{19}\left(y - \dfrac{3}{8}\right)$.

8.4.

1. Кривая среднеквадратической регрессии ζ на ν:

$x = g(y) = E(\zeta|\nu = y) = 1 + |y|\,,$

ν на ζ: $y = h(x) = E(\nu|\zeta = x) = 0\,.$

2. Кривая среднеквадратической регрессии ξ на η:

$x = g(y) = E(\xi|\eta = y) = 1 - \dfrac{1}{2}\sqrt{1-y}\,,$

η на ξ:

$y = h(x) = E(\eta|\xi = x) = \dfrac{1}{2}(2x - x^2)\,.$

3. Кривая среднеквадратической регрессии ξ на η:

$$x = g(y) = E(\xi|\eta = y) = \frac{1}{3}\frac{(1-y)(1+2y)}{1+y}\,,$$

η на ξ:

$$y = h(x) = E(\eta|\xi = x) = \frac{1}{2}(1-x)\,.$$

9.3

1. $E\xi = k\theta\,,\ D\xi = k\theta^2\,.$

2. a) $\xi = \begin{cases} -2\,, & \text{с вероятностью } \dfrac{1}{4}\,, \\ 0\,, & \text{с вероятностью } \dfrac{1}{2}\,, \\ 2\,, & \text{с вероятностью } \dfrac{1}{4}\,, \end{cases}$ b) $\xi \sim U\{-2,-1,1,2\}\,,$

c) $\xi = \begin{cases} -3\,, & \text{с вероятностью } \dfrac{1}{6}\,, \\ 0\,, & \text{с вероятностью } \dfrac{2}{3}\,, \\ 3\,, & \text{с вероятностью } \dfrac{1}{6}\,. \end{cases}$

3. $\varphi_\xi(t) = \begin{cases} \dfrac{2a^2}{t^2}\left(1-\cos\dfrac{t}{a}\right)\,, & t \neq 0\,, \\ 1\,, & t = 0\,. \end{cases}$

4. $\varphi_\xi(t) = \begin{cases} e^{it}\left(\dfrac{2}{it}+\dfrac{2}{t^2}\right) - \dfrac{2}{t^2}\,, & t \neq 0\,, \\ 1\,, & t = 0\,, \end{cases} \qquad E\xi = \dfrac{2}{3}\,.$

5. $f_\xi(x) = \dfrac{1}{2}e^{-|x|}\,, E\xi = 0\,,\ D\xi = 2\,.$

6. a) $\varphi_{\xi_k}(t) = pe^{it} + 1 - p\,,$ b) $\varphi_{\eta_n}(t) = (pe^{it} + 1 - pf)^n\,,$

c) Используя разложение бинома Ньютона, получаем

$\varphi_{\eta_n}(t) = \sum\limits_{k=0}^{n} C_n^k p^k (1-p)^{n-k} e^{ikt}$, следовательно

$P(\eta_n = k) = C_n^k p^k (1-p)^{n-k}$.

C.5

1. a) $E\vec{\xi} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\Lambda = \begin{pmatrix} 2 & 1 \\ 1 & 5 \end{pmatrix}$,

b) $f_{\xi_1}(x) = \dfrac{1}{\sqrt{4\pi}} e^{-\frac{(x-1)^2}{4}}$, $f_{\xi_2}(x) = \dfrac{1}{\sqrt{10\pi}} e^{-\frac{x^2}{10}}$,

c) $P\big(2\xi_1 - 3\xi_2 < -4\big) \approx 0,17$,

d) $P\big(\xi_1 - 1 < \xi_2 < -3\xi_1 + 3\big) = \dfrac{1}{2\pi} \operatorname{arctg} 12 \approx 0,2368$,

e) $y = E(\xi_2|\xi_1 = x) = \dfrac{x-1}{2}$, $x = E(\xi_1|\xi_2 = y) = 1 + \dfrac{y}{5}$.

2. a) $E\vec{\xi} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $\Lambda = \begin{pmatrix} 1 & -1 \\ -1 & 4 \end{pmatrix}$,

b) $f_{\xi_1}(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{(x-1)^2}{2}}$, $f_{\xi_2}(x) = \dfrac{1}{\sqrt{8\pi}} e^{-\frac{x^2}{8}}$,

c) $P\big(-\xi_1 + 3\xi_2 < 2\big) \approx 0,82$,

d) $P\big(2\xi_1 - 3 < \xi_2 < 3\xi_1 - 4\big) = \dfrac{1}{2\pi} \operatorname{arctg} \dfrac{1}{5\sqrt{3}} \approx 0,0183$,

e) $y = E(\xi_2|\xi_1 = x) = -x$, $x = E(\xi_1|\xi_2 = y) = 1 - \dfrac{y+1}{4}$.

Учебное издание

Альшанский Максим Алексеевич

Теория вероятностей и математическая статистика

Корректор *З. Р. Бухонова*
Верстка (LaTeX) *М. А. Альшанского*

Подписано в печать 09.11.2021. Формат 70×100/16.
Бумага писчая. Цифровая печать. Усл. печ. л. 18,1.
Уч.-изд. л. 11,3. Тираж 30 экз. Заказ 255.

Издательство Уральского университета
Редакционно-издательский отдел ИПЦ УрФУ
620049, Екатеринбург, ул. С. Ковалевской, 5
Тел. 8 (343) 375-48-25, 375-46-85, 374-19-41
E-mail rio@urfu.ru

Отпечатано в Издательско-полиграфическом центре УрФУ
620075, Екатеринбург, ул. Тургенева, 4
Тел.: 8 (343) 350-56-64, 350-90-13
Факс: 8 (343) 358-93-06
E-mail: press-urfu@mail.ru

ЛИДЕР УНИВЕРСИТЕТСКОГО КНИГОИЗДАНИЯ
УНИВЕРСИТЕТСКАЯ КНИГА
2020

9 785799 633851